(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 April 2003 (03.04.2003)

PCT

(10) International Publication Number
WO 03/027299 A2

(51) International Patent Classification[7]: C12N 15/82

(21) International Application Number: PCT/EP02/10364

(22) International Filing Date: 16 September 2002 (16.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01870200.1 17 September 2001 (17.09.2001) EP

(71) Applicant (*for all designated States except US*): **CROPDESIGN NV** [BE/BE]; Technologiepark 3, B-9052 Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **DE VEYLDER, Lieven** [BE/BE]; Josef Boddaertstraat 23, B-9031 Drongen (BE). **DE PINHO BAROCCO, Rosa, Maria** [PT/BE]; Fayte 6, B-9661 Parike (BE). **MIRONOV, Vladimir** [BE/BE]; Gentbruggestraat 142, B-9040 St-Amandsberg (BE). **INZE, Dirk** [BE/BE]; Driesstraat 18, B-9310 Moorsel-Aalst (BE).

(74) Agents: **DE CLERCQ, Ann** et al.; De Clercq, Brants & Partners CV, E. Gevaertdreef 10a, B-9830 St-Martens-Latem (BE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/027299 A2

(54) Title: A PLANT CYCLIN DEPENDENT KINASE-LIKE PROTEIN, ITS INTERACTORS AND USES THEREOF

(57) Abstract: The present invention relates to methods for modifying plant growth and development processes comprising modulating expression of a plant cyclin dependent kinase-like gene and/or one of its interacting proteins or homologues, derivatives or fragments thereof. The invention further relates to the use of vectors for performing the present invention and to transgenic plants produced therewith having altered plant growth and development characteristics compared to their isogenic counterparts. Preferably, the characteristics modified by the present invention include growth rate, yield, senescence, flowering and photosynthesis.

A PLANT CYCLIN DEPENDENT KINASE- LIKE PROTEIN, ITS INTERACTORS AND USES THEREOF

FIELD OF THE INVENTION

The present invention relates to methods for modifying plant growth and development processes comprising modulating expression of a plant cyclin dependent kinase-like gene and/or one of its interacting proteins or derivatives thereof. The invention further relates to vectors useful for performing the present invention and to transgenic plants produced therewith having altered plant growth and development characteristics compared to their isogenic counterparts. Preferably, the characteristics modified by the present invention include growth rate, yield, senescence, flowering and photosynthesis.

BACKGROUND TO THE INVENTION

Dividing eukaryotic cells go through a highly ordered sequence of events termed the cell cycle (Morgan, 1997). The basic mechanisms controlling the progression through the different steps of the cell cycle appear to be conserved in all higher eukaryotes. Transitions through and between the different stages of the mitotic cell cycle depend on the activity of a complex consisting of a cyclin-dependent kinase (CDK) and a specific subset of cyclins. Cyclins target the kinase activity of CDKs to specific substrates. The association of CDKs with various cyclins allows for the formation of multiple protein kinase complexes with specialized cell-cycle functions. Additional factors that regulate CDK activity include CDK inhibitors, CDK activating kinase and CDK phosphatase. Eukaryote genomes typically encode multiple CDK and CDK-like genes. International patent application WO 00/56905 generally describes a method for modifying various plant characteristics by expression of at least two cell cycle interacting proteins. The patent application for instance mentions co-expression of CDKs and their interacting cyclins. Considerable progress has been made in the characterization of CDK and cyclin proteins that play a role in cell cycle progression in yeast, animal systems, and also in plants. For example, in *Arabidopsis thaliana*, two CDKs have been identified as major regulators of the cell cycle (Mironov, De Veylder et al., 1999). These CDKs have recently been renamed Arath;CDKA;1 and Arath;CDKB;1 and represent two major plant CDK groups, CDKA and CDKB (Joubes, Chevalier et al., 2000). The CDKA-type proteins contain the characteristic PSTAIRE motif and seem to be involved in cell proliferation or maintenance of cell division competence in non-proliferating tissues.

Members of the CDKB group play a role in mitosis and contain the PPTA/TTLRE motif, which is unique to plants.

A small group of CDK-like proteins have been identified in plants that are characterized by the presence of the PITAIRE motif in the cyclin binding domain (Joubes, Chevalier, et al., 2000). It was proposed that these PITAIRE CDK's be named CDKC. It has been suggested that CDKC-type kinases are not directly involved in cell cycle control, although their function is unknown (Mironov, De Veylder, et al., 1999). Therefore, one of the objects of the present invention is to identify the protein interactors of a CDKC-type protein and their biological functions. Modulating expression of these proteins allows manipulating the biological processes that they control. It is a further object of this invention to modulate these biological processes which are particularly useful for applications in agriculture. The invention provides a solution to at least several of the objects above by providing any of the methods described herein.

SUMMARY OF THE INVENTION

In the present invention, the protein interactors of the *Arabidopsis thaliana* Arath;CDKC;2 protein are disclosed. These proteins include the cyclin regulator of this CDKC as well as targets or additional protein subunits of the CDKC/cyclin complex, including DNA/RNA binding proteins and proteins involved in photosynthesis and chloroplast development and/or function.

The present invention generally relates to a method for modifying plant biochemical and physiological characteristics, such as one or more developmental and/or environmental processes, including but not limited to the modification of plastid development, and/or photosynthetic capacity and greening, and/or stress-induced responses, and/or timing of senescence, and/or timing of flowering, and/or seed development, and/or seed yield, said method comprising expressing a CDKC-type protein or a mutant form thereof alone or in combination with one of its interacting partners, in the plant, operably under the control of a regulatable promoter, preferably a cell- or tissue- or organ-specific promoter. The present invention extends to the use of genetic constructs for performing the methods of the invention and to transgenic plants produced therewith having altered growth and/or development and/or physiological characteristics compared to their otherwise isogenic plants.

DETAILED DESCRIPTION OF THE INVENTION

In order to manage problems related to plant growth and yield, it is of utmost importance not only to isolate plant genes but especially to characterize the function of the encoded proteins. Only when the function of a protein or gene is known, can it be rationally applied towards influencing the growth and yield of the plant as a whole.

According to a first embodiment the present invention relates to a method for altering or modifying biochemical and physiological characteristics of a plant or plant cell such as developmental and/or growth and/or yield characteristics comprising modulating the expression in a plant or plant cell of at least one first nucleic acid encoding a plant CDKC kinase, a homologue or a derivative thereof or an enzymatically active fragment thereof and/or at least one second nucleic acid encoding a CDKC kinase interacting protein, a homologue or a derivative thereof or an enzymatically active fragment thereof

The expression "modulating" or "altering" the expression relates to methods for altering the expression of at least one first and/or a second nucleic acid in specific cells or tissues.

In the context of the present invention the term "modulation" or "altering" relates to enhancing or decreasing the expression or, alternatively may relate to upregulating or downregulating the expression. According to at least one preferred embodiment of the invention, downregulated or decreased expression of said nucleic acid is envisaged.

According to the invention, the "nucleic acid" may be the wild type endogenous nucleic acid whose expression is modulated or may be a paralogue or orthologue, i.e. a homologous nucleic acid derived from the same or another species.

The present invention involves the modulation of expression of at least one nucleic acid encoding a plant CDKC kinase. The current classification of plant CDKs is based mainly on sequence similarity and this organization corresponds well with differential functions of each CDK class. Unlike members of the class A and B cyclin dependent kinases, type C CDK kinases are thought not to be directly involved in cell cycle regulation. However, their precise function and their cyclin partner(s) or other protein interactors were hitherto unknown.

The cyclin dependent kinase-like proteins of the present invention specifically belong to the 'PITAIRE cluster/CDKC Plants' group as illustrated in Figure 2 of Joubes et al., 2000. This group contains cyclin dependent kinase-like proteins from plant and animal origin that can be differentiated from other cyclin dependent kinases based on

comparative amino acid sequence analysis as described (Joubes, Chevalier et al., 2000). More specifically, the cyclin dependent kinase-like proteins of the present invention belong to the CDKC plant cyclin dependent kinases in this group that are characterized by the presence of the PITAIRE motif in their cyclin binding box. These plant specific cyclin dependent kinase-like proteins are therefore also termed the 'PITAIRE kinases' in the present invention. It was proposed to group the plant PITAIRE kinases in the *CDKC* class to differentiate them from other cyclin dependent kinases (Joubes, Chevalier et al., 2000). The *CDKC* class currently contains four different CDKs from three plant species but it is envisaged that other plant species have similar, still unidentified, CDKs as well and these also fall within the scope of this invention. New members of this proposed CDKC class may or may not contain the identical PITAIRE motif.

In the present invention, a two-hybrid screen was performed to identify and isolate gene products interacting with Arath;CDKC;2 which belongs to the C class of cyclin dependent kinases.

According to preferred embodiments, the invention thus relates to any of the methods of the invention wherein said plant CDKC kinase is the Arath;CDKC;2 represented by SEQ ID NO 2, or a homologue, derivative or an enzymatically active fragment thereof.

In the Examples section, methods are described how to identify "CDKC kinase interacting proteins". Several protein interacting partners have been identified and are described herein but other CDKC interacting proteins still have to be identified using the same strategy as herein described. Furthermore, similar two-hybrid screenings can be performed using other members of the type C CDK kinase family. It should be clear that the invention thus also relates to the use of said proteins in the methods of the invention.

A first protein identified in the two-hybrid screen is an *Arabidopsis* protein which was designated CYCT1At for cyclin T1-like protein from *Arabidopsis thaliana* (represented by SEQ ID NOs 3 and 4). It is clearly demonstrated in the Examples section of the present invention that CYCT1At specifically interacts with Arath;CDKC;2 but not with a member of the CDKA or CDKB class of CDKs.

The plant cyclin T1-like proteins of the present invention are defined as cyclin-like proteins of plant origin that specifically bind to the plant CDKC kinases to form a heterodimer complex. Such cyclin T1-like protein/CDKC heterodimers may be active in

phosphorylating proteins and may contain additional proteins to form a dynamic multiprotein complex.

Therefore, a major embodiment of the current invention relates to the specific and functional association between a member of the class C cyclin dependent kinases, Arath;CDKC;2, and the cyclin CYCT1At. A further embodiment of the present invention thus relates to the identification and characterization of a novel plant CDK/cyclin complex.

In a most preferred embodiment the invention relates to any of the methods described herein wherein said plant CDKC kinase is represented by SEQ ID NO 2 and wherein said CDKC kinase interacting protein is CYCT1At represented by SEQ ID NO 4, or a homologue thereof.

Furthermore, in plants also other CDK/cyclin complexes could exist that are structurally and functionally related to Arath;CDKC;2/ CYCT1At complex. It should be understood that these also fall within the scope of this invention.

To the scope of the current invention also belong plant polypeptides which have, compared to the CYCT1At protein, similar properties in that they specifically bind to a member of the C class of plant cyclin dependent kinases such as Arath;CDKC;2.

The present invention also relates to Arath;CDKC;2 interactors identified and characterised herein that are different from CYCT1At. Several proteins have been identified in the present invention that may either be a target or substrate of the Arath;CDKC;2 protein or of the Arath;CDKC;2/CYCT1At complex or that may be a part of a multiprotein complex that includes Arath;CDKC;2 and/or CYCT1At. Furthermore, the identification of additional interactors of Arath;CDKC;2 has provided additional information on the function(s) of this kinase in plant cells and provides new ways to manipulate these function(s).

Several other protein interactors of Arath;CDKC;2 have been identified and isolated as described herein that are either transcription factors or proteins involved in nuclear processes. These proteins include the DNA binding protein AtGT1 (represented by SEQ ID NOs 15 to 17), a ribonucleoprotein (RNP; represented by SEQ ID NOs 10 and 11), and a protein designated herein as AtCDKCIP1 for *Arabidopsis thaliana* CDKC interacting protein 1 (represented by SEQ ID NOs 12 to 14) that is a putative transcription factor as disclosed herein (see Example 4).

The terms "ribonucleoproteins" or "RNPs" refer to very abundant RNA-binding proteins that play an important role in the metabolism of pre-mRNA, bind pre-mRNAs attached

to RNA polymerase II elongation complexes, and influence pre-mRNA maturation at different levels, such as alternative splicing and mRNA export. The interaction of Arath;CDKC;2 with RNP might be essential for the regulation of diverse processing events, including mRNA splicing and transport.

The "AtGT-1" protein relates to a plant transcription factor identified by its specific binding activity to promoters of light-regulated genes. The interaction of Arath;CDKC;2 with GT-1 suggests an involvement of Arath;CDKC;2 in light-regulated transcription. These findings indicated that Arath;CDKC;2 in functional association with a cyclin interactor and/or one or more other protein interactors, is not directly involved in cell cycle regulation but instead plays a role in nuclear processes such as transcription regulation and/or RNA processing events.

Therefore, according to a further embodiment, the present invention relates to a method for altering developmental and/or growth and/or yield characteristics of a plant or plant cell said method comprising modulating transcription regulation.

Still other Arath;CDKC;2 interacting protein partners were identified that play a role in photosynthesis and chloroplast development, including ribulose bisphosphate carboxylase (rubisco) activase (represented by SEQ ID NOs 5 and 6) and the DAG-like protein (represented by SEQ ID NOs 7 to 9).

Rubisco activase controls the process of photosynthesis by making the activity of rubisco responsive to light intensity. The term "DAG-like protein" refers to proteins whose expression is required for the expression of nuclear genes that encode proteins implicated in light-regulated gene expression such as the chlorophyl a/b binding protein (CAB) and rubisco. DAG has been shown to be targeted to the plastids. However, the present work indicates that DAG proteins may also directly interact with nuclear proteins such as CDKC;2, being targeted to the nucleus where it may interact with the transcription machinery. The evidence provided herein that rubisco activase and a DAG-like protein interact with Arath;CDKC;2 indicates that this kinase, probably in association with its cyclin binding partner, is a regulator of proteins involved in plastid development.

Another aspect of this invention is the characteristic expression pattern of the Arath;CDKC;2 gene and the CYCT1At as determined by real-time PCR (Example 5) and *in situ* hybridization (Example 6). These results showed that Arath;CDKC;2 and CYCT1At transcripts are present in seedlings, root tissue, rosettes, stems and flowers but were most abundant in flower tissue. Arath;CDKC;2 transcripts are expressed in a

tissue specific and developmentally regulated fashion. Most importantly, the *in situ* hybridization experiments demonstrated that Arath;CDKC;2 transcripts are present in terminally differentiated tissues but not in actively dividing tissues such as meristems, which confirmed the results of the two hybrid screening in that this kinase is not directly involved in cell division control but plays a role in differentiated tissues such as flowers.

Therefore, according to a still further embodiment, the present invention relates to a method for altering developmental and/or growth and/or yield characteristics of a plant or plant cell said method comprising modulating photosynthesis and/or chloroplast development. Alternatively, the invention relates to a method for enhancing the photosynthetic capacity of a plant or plant cell comprising modulating the expression in a plant or plant cell of at least one first nucleic acid encoding a plant CDKC kinase, a homologue or a derivative thereof or an enzymatically active fragment thereof and/or at least one second nucleic acid encoding a CDKC kinase interacting protein, a homologue or a derivative thereof or an enzymatically active fragment thereof.

As such, it can be summarized that in the present invention Arath;CDKC;2 interacting proteins were identified that play a role in transcription regulation and/or photosynthesis and/or chloroplast development. Modulating the expression level or activity of the Arath;CDKC;2 protein in a plant or plant cell, either by itself or in combination with modulated expression of one or more of its protein interactors selected from the list of CYCT1At, AtGT1, a ribonucleoprotein, AtCDKCIP1, the DAG-like protein or rubisco activase, can be used to modulate the growth and development characteristics of a plant including but not limited to chloroplast development and photosynthesis.

All together these data lead to the consideration that the plant CDKC/cyclin T complex is not involved in cell cycle control but rather interacts with specific components of transcriptional machinery to repress chloroplast development in flower epidermal cells.

One more preferred embodiment thus relates to a method as described above resulting in an increase in the number of flowers and/or seeds and/or fruits of a plant.

In yet another preferred embodiment, the invention relates to any of the methods of the invention wherein a plant CDKC kinase represented by SEQ ID NO 2 or encoded by SEQ ID NO 1 is used, and wherein said CDKC kinase interacting protein is chosen from the polypeptides represented by any of SEQ ID NOs 4, 6, 8, 9, 11, 13, 14, 16 or 17 or encoded by any of SEQ ID NOs 3, 5, 7, 10, 12 or 15.

One way of modulating the expression of a CDKC kinase or a CDKC kinase interacting protein comprises the stable integration in an expressible form into the genome of a

plant or in specific plant cells or tissues of said plant of at least one first nucleic acid encoding said CDKC kinase, a homologue or a derivative thereof or an enzymatically active fragment thereof and /or at least one second nucleic acid encoding said CDKC kinase interacting protein, a homologue or a derivative thereof or an enzymatically active fragment thereof.

The term "expressible form" should be understood as containing the control sequence needed for expression. One way of expression according to the invention relates to "ectopic expression" or "ectopic overexpression" of a gene or a protein which refers to expression patterns and/or expression levels of said gene or protein normally not occurring under natural conditions. Ectopic expression can be achieved in a number of ways including operably linking of a coding sequence encoding said protein to an isolated homologous or heterologous promoter in order to create a chimeric gene and/or operably linking said coding sequence to its own isolated promoter (i.e. the unisolated promoter naturally driving expression of said protein) in order to create a recombinant gene duplication or gene multiplication effect.

According to another preferred embodiment the invention relates to any of the above methods wherein downregulation of expression of said first or second nucleic acid is achieved. Preferably, said method comprising the stable integration into the genome of a said plant or said plant cells of at least one nucleic acid causing downregulation of said first or second nucleic acids.

Methods for downregulation of expression of endogenous genes are well known in the art and may comprise the use of sense or antisense copies of at least part of the endogenous gene in the form of direct or inverted repeats.

Therefore, the invention also relates to the above method wherein the nucleic acid causing downregulation comprises at least part of an antisense version of said first or second nucleic acid.

The term "antisense version" relates to a nucleic acid which is the "antisense" of said nucleic acid and which is able to hybridise therewith. It should be clear that "at least part" of said nucleic acid may suffice to achieve the desired result.

In case of integrating an extra copy of a sense version of a CDKC kinase or CDKC kinase interacting protein, downregulation of expression can also be obtained: the introduced gene suppresses its own expression and that of the homologous genes, through a phenomenon termed cosuppression, well known to those skilled in the art.

Preferably, in the methods for downregulation of expression of CDKC kinases or CDKC kinase interacting proteins, said first nucleic acid is represented by SEQ ID NO 1 and said second nucleic acid is chosen from the group of nucleic acids represented in SEQ ID NOs 3, 5, 7, 10, 12 or 15.

In a more specific embodiment, the invention relates to the methods as described above wherein a nucleic acid encoding the CYCT1At protein, represented by SEQ ID NO 4, or a homologue thereof is downregulated. The CYCT1At protein is shown herein to be the cyclin partner of the Arath;CDKC;2 kinase.

According to a most specific embodiment, the present invention relates to any of the methods of the invention wherein said plant CDKC kinase is represented by SEQ ID NO 2 or a derivative thereof or an enzymatically active fragment thereof and wherein said CDKC kinase interacting protein is CYCT1At represented by SEQ ID NO 4 or a derivative thereof or an enzymatically active fragment thereof.

The present invention also relates to methods for the production of a transgenic plant having altered growth and/or yield characteristics comprising:

- transforming a plant or a plant cell with a DNA construct comprising a gene promoter sequence, preferably a tissue- or cell-specific promoter, with (i) at least one open reading frame encoding at least one functional portion of a CDKC kinase, a homologue or a derivative thereof, preferably a CDKC kinase encoded by a nucleic acid represented by SEQ ID NO 2, , and/or (ii) at least one second open reading frame encoding at least one functional portion of a CDKC kinase interacting protein, a homologue or a derivative thereof, preferably a CDKC kinase interacting protein represented by any of SEQ ID NOs 4, 6, 8, 9, 11, 13, 14, 16 or 17, to provide a transgenic cell;
- providing means for altering the expression of said nucleic acid, preferably by gene silencing; and
- cultivating the transgenic cell under conditions promoting regeneration and mature plant growth.

The expression "a functional portion" relates to a nucleic acid encoding an enzymatically active fragment of a CDKC kinase or CDKC kinase interacting protein. The expression " a functional portion" also relates to a nucleic acid corresponding to a sense or antisense fragment or version of a CDKC kinase or CDKC kinase interacting protein which can be used in any of the methods for downregulation of expression of its endogenous counterpart. It should be clear that such sense or antisense fragments do

not necessarily need to encode the CDKC kinase or CDKC kinase interacting protein or an enzymatically active fragment thereof.

The invention further relates to a method for the production of a transgenic plant having altered growth and/or yield characteristics comprising:

- transforming a plant or a plant cell with a DNA construct comprising at least one nucleic acid as defined in any of the methods relating to the downregulation of expression of a CDKC kinase or CDKC kinase interacting protein, under the control of a promoter sequence, preferably a cell- or tissue specific promoter, to provide a transgenic cell; and
- cultivating the transgenic cell under conditions promoting regeneration and mature plant growth.

Also according to the invention are the methods herein described comprising the use of promoters which are not cell- or tissue-specific but which are constitutive promoters. In tables A and B, examples are given of such cell- and tissue-specific promoters and constitutive promoters.

The plant cells or plants used in the methods of the present invention include all plants or cells of plants which belong to the superfamily *Viridiplantae*, including both monocotyledonous and dicotyledonous plants. Two of the most preferred plants for use in the methods of the invention are *Arabidopsis thaliana* and *Oryza sativa* (rice) or plant cells or tissues derived thereof.

The invention also relates to any transgenic plant obtainable by any of the methods described herein.

According to yet another embodiment the invention relates to a method for identifying and obtaining compounds that interfere with the interaction between a CDKC kinase and a CDKC kinase interacting protein comprising the steps of :

(a) providing an expression system wherein a CDKC kinase, a homologue or a derivative or a fragment thereof, and a CDKC kinase interacting protein, a homologue, a derivative or a fragment thereof are expressed, preferably said CDKC kinase is represented by SEQ ID NO 2 and said CDKC kinase interacting protein is represented by any of SEQ ID NOs 4, 6, 8, 9, 11, 13, 14, 16 or 17,

(b) interacting at least one compound with the complex formed by the expressed polypeptides as defined in (a), and,

(c) measuring the effect of said compound on the binding between the interacting proteins as defined in (a) or measuring the activity of said complex;

(d) optionally identifying said compound.

In a preferred embodiment, the invention relates to the above compound screening method wherein said compound inhibits the activity of said protein complex or inhibits the formation of a complex between said proteins. In an alternative embodiment, the invention relates to the above compound screening method wherein said compound enhances the activity of said protein complex or promotes the formation of a complex between said proteins or influences the activity of said complex.

The invention relates to any compound obtainable by any of the compound screening methods described.

The use of said compounds identified by means of any of said method as a plant growth regulator or as a plant herbicide is also part of the present ivention.

The invention further relates to a method for the production of a plant growth regulator or herbicide composition comprising the steps of any of the compound screening methods and formulating the compounds obtained from said steps in a suitable form for the application in agriculture or plant cell or tissue culture.

The invention also relates to a method for the design of or screening for growth-promoting chemicals or herbicides comprising the use of a nucleic acid encoding a CDKC kinase, a homologue or a derivative or a fragment thereof, and a CDKC kinase interacting protein, a homologue, a derivative or a fragment thereof.

According to a more general embodiment the invention relates the use of a nucleic acid encoding CDKC kinase, a homologue or a derivative or a fragment thereof, and a CDKC kinase interacting protein, a homologue, a derivative or a fragment thereof for modulating transcription regulation processes or for enhancing the photosynthetic capacity of specific plants.

According to more specific embodiments the invention further relates to the use of a nucleic acid encoding a CDKC kinase, a homologue or a derivative or a fragment thereof, and a CDKC kinase interacting protein, a homologue, a derivative or a fragment thereof for increasing yield, stimulating growth or for increasing the number of flowers and/or seeds and/or fruits per plant.

Definitions and elaborations to the embodiments

Those skilled in the art will be aware that the invention described herein is subject to variations and modifications other than those specifically described. It is to be understood that the invention described herein includes all such variations and modifications. The invention also includes all such steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any or more of said steps or features.

Nucleic acids are written left to right in 5' to 3' orientation, unless otherwise indicated; amino acid sequences are written left to right in amino to carboxy orientation. Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides may be referred to by their commonly accepted single-letter codes.

Numeric ranges are inclusive of the numbers defining the range.

The term 'gene(s)', 'polynucleotide', 'nucleic acid', 'nucleotide sequence', 'nucleic acid ' or 'nucleic acid molecule(s)' as used herein refers to a polymeric form of a deoxyribonucleotides or ribonucleotide polymer of any length, either double- or single-stranded, or analogs thereof, that have the essential characteristic of a natural ribonucleotide in that they can hybridize to nucleic acids in a manner similar to naturally occurring polynucleotides. A great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those skilled in the art. For example, methylation, 'caps' and substitution of one or more of the naturally occurring nucleotides with an analog. Said terms also include peptide nucleic acids. The term polynucleotide as used herein includes such chemically, enzymatyically or metabolically modified forms of polynucleotides. 'Sense strand' refers to a DNA strand that is homologous to a mRNA transcript thereof, 'antisense strand' refers to the complementary strand of the sense strand.

By 'encoding' or 'encodes' with respect to a specified nucleotide sequence is meant comprising the information for translation into a specified protein. A nucleic acid encoding a protein may contain non-translated sequences such as 5' and 3' untranslated regions (5' and 3' UTR) and introns or it may lack intron sequences such as for example in cDNAs. An 'open reading frame' or '(ORF)' is defined as a nucleotide sequence that encodes a polypeptide. The information by which a protein is encoded is specified by the use of codons. Typically, the amino acid sequence is

encoded by the nucleic acid using the 'universal' genetic code but variants of this universal code exist (see for example Proc. Natl. Acad. Sci. U.S.A 82: 2306-2309 (1985)). The boundaries of the coding sequence are determined by a translation start codon at the 5'end and a translation stop codon at the 3'-terminus. As used herein 'full-length sequence' with respect to a specific nucleic acid or its encoded protein means having the entire amino acid sequence of a native protein. In the present invention, comparison to known full-length homologous (orthologous or paralogous) sequences is used to identify full-length sequences. Also, for a mRNA or cDNA, consensus sequences present at the 5' and 3' untranslated regions aid in the identification of a polynucleotide as full-length. For a protein, the presence of a start- and stopcodon aid in identifying the polypeptide as full-length. When the nucleic acid is to be expressed, advantage can be taken of known codon preferences or GC content preferences of the intended host as these preferences have been shown to differ (see e.g. http://www.kazusa.or.jp/codon/; Murray et al., Nucl. Acids Res. 17: 477-498 (1989)). Because of the degeneracy of the genetic code, a large number of nucleic acids can encode any given protein. As such, substantially divergent nucleic acid sequences can be designed to effect expression of essentially the same protein in different hosts. Conversely, genes and coding sequences essentially encoding the same protein isolated from different sources can consist of substantially different nucleic acid sequences.

The term 'control sequence' or 'regulatory sequence' or 'regulatory element' refers to regulatory nucleic acid sequences which are necessary to effect the expression of sequences to which they are ligated. The control sequences differ depending upon the intended host organism and upon the nature of the sequence to be expressed. For expression of a protein, in prokaryotes, the control sequences generally include a promoter, a ribosomal binding site, and a terminator. In eukaryotes, control sequences generally include promoters, terminators and, in some instances, enhancers, introns, and/or 5' and 3' untranslated sequences. The term 'control sequence' is intended to include, at a minimum, all components necessary for expression, and may also include additional advantageous components.

As used herein, a 'promoter' includes reference to a region of DNA upstream from the transcription start and involved in binding RNA polymerase and other proteins to start transcription. Reference herein to a 'promoter' is to be taken in its broadest context and includes the transcriptional regulatory sequences derived from a classical eukaryotic genomic gene, including the TATA box which is required for accurate transcription

initiation, with or without a CCAAT box sequence and additional regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or external stimuli, or in a tissue-specific manner. The term 'promoter' also includes the transcriptional regulatory sequences of a classical prokaryotic gene, in which case it may include a -35 box sequence and/or a -10 box transcriptional regulatory sequences. The term 'promoter' is also used to describe a synthetic or fusion molecule, or derivative which confers, activates or enhances expression of a nucleic acid molecule in a cell, tissue or organ. A 'plant promoter' is a promoter capable of initiating transcription in plant cells. 'Tissue-preferred promoters' as used herein refers to promoters that preferentially initiate transcription in certain tissues such as for example in leaves, roots, etc. Promoters which initiate transcription only in certain tissues are referred herein as 'tissue-specific'. Those skilled in the art will be aware that 'inducible promoters' have induced or increased transcription initiation in response to a developmental, chemical, environmental, or physical stimulus and that a 'constitutive promoter' is transcriptionally active during most, but not necessarily all phases of its growth and development. Examples of plant tissue-specific or tissue-preferred promoters are given in Table 1. Examples of constitutive plant promoters are given in Table 2. The term 'terminator' as used herein is an example of a 'control sequence' and refers to a DNA sequence at the end of a transcriptional unit which signals 3'processing and polyadenylation of a primary transcript and termination of transcription. Terminators comprise 3'-untranslated sequences with polyadenylation signals, which facilitate 3'processing and the addition of polyadenylate sequences to the 3'-end of a primary transcript. Terminators active in cells derived from viruses, yeasts, moulds, bacteria, insects, birds, mammals and plants are known and described in the literature. They may be isolated from bacteria, fungi, viruses, animals and/or plants. Additional regulatory elements may include transcriptional as well as translational enhancers. A plant translational enhancer often used is the CaMV omega sequences. The inclusion of an intron has been shown to increase expression levels by up to 100-fold in certain plants (Mait, Transgenic Research 6 (1997), 143-156; Ni, Plant Journal 7 (1995), 661-676).

TABLE 1. Exemplary plant tissue-specific or tissue-preferred promoters

GENE SOURCE	**EXPRESSION PATTERN**	**REFERENCE**
α-amylase (*Amy32b*)	Aleurone	Lanahan, MB, *et al.*, Plant Cell 4: 203-211, 1992; Skriver, K, *et al.*, Proc. Natl. Acad. Sci. (USA) 88: 7266-7270, 1991.
Cathepsin β-like gene	Aleurone	Cejudo, FJ, *et al.*, Plant Mol. Biol. 20: 849-856, 1992.
Agrobacterium rhizogenes rolB	Cambium	Nilsson *et al.*, Physiol. Plant. 100: 456-462, 1997.
PRP genes	cell wall	http://salus.medium.edu/mmg/tierney/html
Chalcone synthase (chsA)	Flowers	Van der Meer *et al.*, Plant Mol. Biol. 15: 95-109, 1990.
LAT52	Anther	Twell *et al.*, Mol. Gen. Genet. 217: 240-245, 1989.
Apetala-3	Flowers	
Chitinase	fruit (berries, grapes, etc)	Thomas *et al.*, CSIRO Plant Industry, Urrbrae, South Australia, Australia; http://winetitles.com.au/gwrdc/csh95-1.html
Rbcs-3A	green tissue (eg leaf)	Lam *et al.*, The Plant Cell 2: 857-866, 1990; Tucker *et al.*, Plant Physiol. 113: 1303-1308, 1992.
Leaf-specific genes	Leaf	Baszczynski *et al.*, Nucl. Acids Res. 16: 4732, 1988.
Chlorella virus adenine methyltransferase gene promoter	Leaf	Mitra and Higgins, Plant Mol. Biol. 26: 85-93, 1994.
AldP gene promoter from rice	Leaf	Kagaya *et al.*, Mol. and Gen. Genet. 248: 668-674, 1995.
Rbcs promoter from rice or tomato	Leaf	Kyozuka et al., Plant Physiol. 102: 991-1000, 1993.
Pinus cab-6	Leaf	Yamamoto *et al.*, Plant Cell Physiol. 35: 773-778, 1994.
Rubisco promoter	Leaf	
Cab (chlorophyll a/b binding protein)	Leaf	
SAM22	senescent leaf	Crowell *et al.*, Plant Mol. Biol. 18: 459-466, 1992.
Ltp gene (lipid transfer gene)		Fleming *et al.*, Plant J. 2: 855-862,

R. japonicum nif gene	Nodule	United States Patent No. 4, 803, 165
B. japonicum nifH gene	Nodule	United States Patent No. 5, 008, 194
GmENOD40	Nodule	Yang *et al.*, The Plant J. 3: 573-585,
PEP carboxylase (PEPC)	Nodule	Pathirana *et al.*, Plant Mol. Biol. 20: 437-450, 1992.
Leghaemoglobin (Lb)	Nodule	Gordon *et al.*, J. Exp. Bot. 44: 1453-1465, 1993.
Tungro bacilliform virus gene	Phloem	Bhattacharyya-Pakrasi *et al.*, The Plant J. 4: 71-79, 1992.
Sucrose-binding protein gene	plasma membrane	Grimes *et al.*, The Plant Cell 4: 1561-1574, 1992.
Pollen-specific genes	pollen; microspore	Albani *et al.*, Plant Mol. Biol. 15: 605, 1990; Albani *et al.*, Plant Mol. Biol. 16: 501, 1991.
Zm13	Pollen	Guerrero *et al.*, Mol. Gen. Genet. 224: 161-168, 1993.
Apg gene	Microspore	Twell *et al.*, Sex. Plant Reprod. 6: 217-224, 1993.
Maize pollen-specific gene	Pollen	Hamilton *et al.*, Plant Mol. Biol. 18: 211-218, 1992.
Sunflower pollen-expressed gene	Pollen	Baltz *et al.*, The Plant J. 2: 713-721, 1992.
B. napus pollen-specific gene	pollen;anther; tapetum	Arnoldo *et al.*, J. Cell. Biochem., Abstract No. Y101, 204, 1992.
Root-expressible genes	Roots	Tingey *et al.*, EMBO J. 6: 1, 1987.
Tobacco auxin-inducible gene	Root tip	Van der Zaal *et al.*, Plant Mol. Biol. 16, 983, 1991.
β-tubulin	Root	Oppenheimer *et al.*, Gene 63: 87, 1988.
Tobacco root-specific genes	Root	Conkling *et al.*, Plant Physiol. 93: 1203, 1990.
B. napus G1-3b gene	Root	United States Patent No. 5, 401, 836
SbPRP1	Roots	Suzuki *et al.*, Plant Mol. Biol. 21: 109-119, 1993.
AtPRP1; AtPRP3	Roots; root hairs	http://salus.medium.edu/mmg/tierney/html
RD2 gene	root cortex	http://www2.cnsu.edu/ncsu/research
TobRB7 gene	root vasculature	http://www2.cnsu.edu/ncsu/research
AtPRP4	Leaves; flowers; lateral root primordia	http://salus.medium.edu/mmg/tierney/html
Seed-specific genes	Seed	Simon *et al.*, Plant Mol. Biol. 5: 191, 1985; Scofield *et al.*, J. Biol. Chem.

		262: 12202, 1987; Baszczynski *et al.*, Plant Mol. Biol. 14: 633, 1990.
Brazil Nut albumin	Seed	Pearson *et al.*, Plant Mol. Biol. 18: 235-245, 1992.
Legumin	Seed	Ellis *et al.*, Plant Mol. Biol. 10: 203-214, 1988.
Glutelin (rice)	Seed	Takaiwa *et al.*, Mol. Gen. Genet. 208: 15-22, 1986; Takaiwa *et al.*, FEBS Letts. 221: 43-47, 1987.
Zein	Seed	Matzke *et al.*, Plant Mol. Biol., 14: 323-332, 1990.
NapA	Seed	Stalberg *et al.*, Planta 199: 515-519, 1996.
Wheat LMW and HMW glutenin-1	Endosperm	Mol Gen Genet 216: 81-90, 1989; NAR 17: 461-462, 1989
Wheat SPA	Seed	Albani *et al*, Plant Cell, 9: 171-184, 1997.
Wheat α,β,γ-gliadins	Endosperm	EMBO 3: 1409-15, 1984
Barley *Itr1* promoter	Endosperm	
Barley B1, C, D, hordein	Endosperm	Theor Appl Gen 98: 1253-1262, 1999; The Plant J. 4: 343-355, 1993; Mol Gen Genet 250: 750-760, 1996.
Barley DOF	Endosperm	Mena *et al.*, The Plant J. 116: 53-62, 1998.
Blz2	Endosperm	EP99106056.7
Synthetic promoter	Endosperm	Vicente-Carbajosa *et al.*, The Plant J. 13: 629-640, 1998.
Rice prolamin NRP33	Endosperm	Wu *et al.*, Plant Cell Physiol. 39: 885-889, 1998
Rice α-globulin Glb-1	Endosperm	Wu *et al.*, Plant Cell Physiol. 39: 885-889, 1998
Rice OSH1	Embryo	Sato *et al.*, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996.
Rice α-globulin REB/OHP-1	Endosperm	Nakase *et al.*, Plant Mol. Biol. 33: 513-522, 1997.
Rice ADP-glucose PP	Endosperm	Trans. Res. 6: 157-168, 1997.
Maize ESR gene family	Endosperm	The Plant J. 12: 235-246, 1997.
Sorghum γ-kafirin	Endosperm	DeRose RT *et al.*, Plant Mol. Biol. 32: 1029-1035, 1996.
KNOX	Embryo	Postma-Haarsma *et al.*, Plant Mol. Biol. 39:257-271, 1999.
Rice oleosin	Embryo and aleuron	Wu *et al.*, J. Biochem., 123: 386, 1998.
Sunflower oleosin	seed (embryo and dry seed)	Cummins *et al.*, Plant Mol. Biol. 19: 873-876, 1992.

LEAFY	shoot meristem	Weigel *et al.*, Cell 69: 843-859, 1992.
Arabidopsis thaliana knat1	shoot meristem	Accession number AJ131822
Malus domestica kn1	shoot meristem	Accession number Z71981
CLAVATA1	shoot meristem	Accession number AF049870
Stigma-specific genes	Stigma	Nasrallah *et al.*, Proc. Natl. Acad. Sci. USA 85: 5551, 1988; Trick *et al.*, Plant Mol. Biol. 15: 203, 1990.
Class I patatin gene	Tuber	Liu *et al.*, Plant Mol. Biol. 153: 386-395, 1991.
PCNA rice	Meristem	Kosugi *et al.*, Nucl. Acids Res. 19: 1571-1576, 1991; Kosugi S. and Ohashi Y, Plant Cell 9: 1607-1619, 1997.
Pea TubA1 tubulin	Dividing cells	Stotz and Long, Plant Mol. Biol. 41: 601-614, 1999.
Arabidopsis cdc2a	cycling cells	Chung and Parish, FEBS Lett, 362: 215-219, 1995.
Arabidopsis Rop1A	Anthers; mature pollen + pollen tubes	Li *et al.*, Plant Physiol.118: 407-417, 1998.
Arabidopsis AtDMC1	Meiosis-associated	Klimyuk and Jones, The Plant J. 11: 1-14, 1997.
Pea PS-IAA4/5 and PS-IAA6	Auxin-inducible	Wong *et al.*, Plant J. 9: 587-599, 1996.
Pea farnesyltransferase	Meristematic tissues; phloem near growing tissues; light- and sugar-repressed	Zhou *et al.*, Plant J. 12: 921-930, 1997.
Tobacco (*N. sylvestris*) cyclin B1;1	Dividing cells / meristematic tissue	Trehin *et al.*, Plant Mol. Biol. 35: 667-672, 1997.
Catharanthus roseus Mitotic cyclins CYS (A-type) and CYM (B-type)	Dividing cells / meristematic tissue	Ito *et al.*, The Plant J. 11: 983-992, 1997.
Arabidopsis cyc1At (=cyc B1;1) and cyc3aAt (A-type)	Dividing cells / meristematic tissue	Shaul *et al.*, Proc. Natl. Acad. Sci. U.S.A 93: 4868-4872, 1996.
Arabidopsis tef1 promoter box	Dividing cells / meristematic tissue	Regad *et al.*, Mol. Gen. Genet. 248: 703-711, 1995.
Catharanthus roseus cyc07	Dividing cells / meristematic tissue	Ito *et al.*, Plant Mol. Biol. 24: 863-878, 1994.

TABLE 2. Exemplary constitutive plant promoters for use in the performance of the current invention.

GENE SOURCE	REFERENCE
Actin	McElroy *et al.*, Plant Cell 2: 163-171, 1990.
CAMV 35S	Odell *et al.*, Nature 313: 810-812, 1985.
CaMV 19S	Nilsson *et al.*, Physiol. Plant. 100: 456-462, 1997.
GOS2	de Pater *et al.*, The Plant J. 2: 837-44, 1992.
Ubiquitin	Christensen *et al.*, Plant Mol. Biol. 18: 675-689, 1992.
Rice cyclophilin	Buchholz *et al.*, Plant Mol Biol. 25: 837-43, 1994.
Maize H3 histone	Lepetit *et al.*, Mol. Gen. Genet. 231: 276-285, 1992.
Actin 2	An *et al.*, The Plant J. 10: 107-121, 1996.

The term 'operably linked' as used herein refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control sequence 'operably linked' to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences. In case the control sequence is a promoter, it is obvious for a skilled person that double-stranded nucleic acid is used.

In the context of the current invention, 'ectopic expression' or 'ectopic overexpression' of a gene or a protein refers to expression patterns and/or expression levels of said gene or protein normally not occurring under natural conditions. Ectopic expression can be achieved in a number of ways including operably linking of a coding sequence encoding said protein to an isolated homologous or heterologous promoter in order to create a chimeric gene and/or operably linking said coding sequence to its own isolated promoter (i.e. the unisolated promoter naturally driving expression of said protein) in order to create a recombinant gene duplication or gene multiplication effect. With "ectopic co-expression" is meant the ectopic expression or ectopic overexpression of two or more genes or proteins. The same or, more preferably, different promoters are used to confer expression of said genes or proteins.

'Dominant negative version or variant' refers to a mutant protein, which interferes with the activity of the corresponding wild-type protein.

'Downregulation of expression' as used herein means lowering levels of gene expression and/or levels of active gene product and/or levels of gene product activity.

This can be achieved by gene silencing strategies as described by e.g. Angell and Baulcombe 1998 (WO9836083), Lowe et al. 1989 (WO9853083), Lederer et al. 1999 (WO9915682) or Wang et al. 1999 (WO9953050). Genetic constructs aimed at silencing gene expression may have the nucleotide sequence of said gene (or one or more parts thereof) contained therein in a sense and/or antisense orientation relative to the promoter sequence. Another method to downregulate gene expression comprises the use of ribozymes, e.g. as described in Atkins et al. 1994 (WO9400012), Lenee et al. 1995 (WO9503404), Lutziger et al. 2000 (WO0000619), Prinsen et al. 1997 (WO9713865) and Scott et al. 1997 (WO9738116). Still another method to downregulate gene expression comprises e.g. insertion mutagenesis (e.g. T-DNA insertion or transposon insertion).

Immunomodulation is another example of a technique capable of downregulation levels of active gene product and/or of gene product activity and comprises administration of or exposing to or expressing antibodies to said gene product to or in cells, tissues, organs or organisms wherein levels of said gene product and/or gene product activity are to be modulated. Such antibodies comprise "plantibodies", single chain antibodies, IgG antibodies and heavy chain camel antibodies as well as fragments thereof. Modulating, including lowering, the level of active gene products or of gene product activity can furthermore be achieved by administering or exposing cells, tissues, organs or organisms to an inhibitor or activator of said gene product. Such inhibitors or activators include proteins and chemical compounds identified according to the methods of the present invention.

The terms 'protein' and 'polypeptide' are interchangeable used in this application and refer to a polymer of amino acids. These terms do not refer to a specific length of the molecule and thus peptides and oligopeptides are included within the definition of polypeptide. This term also refers to or includes post-translational modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations, sulfations and the like. These modifications are well known to those skilled in the art and examples are described by Wold F., Posttranslational Protein Modifications: Perspectives and Prospects, pp. 1-12 in Posttranslational Covalent Modification of Proteins, B.C. Johnson, Ed., Academic Press, New York (1983) and Seifter et al., Meth. Enzymol. 182: 626-646 (1990). Included within the definition are, for example, polypeptides containing one or more analogues of an amino acid (including, for example, unnatural amino acids, etc.), polypeptides with substituted linkages, as well as other naturally and non-naturally occurring modifications known in the art.

The term 'amino acid'. 'amino acid residue' or 'residue' are used interchangeably herein to refer to an amino acid that is incorporated into a protein, polypeptide, or peptide. The amino acid may be a naturally occurring amino acid and may be a known analogue of natural amino acids that can function in a similar manner as naturally occurring amino acids.

As used herein 'homologues' of a protein of the invention are those peptides, oligopeptides, polypeptides, proteins and enzymes which contain amino acid substitutions, deletions and/or additions relative to said protein, providing similar biological activity as the unmodified polypeptide from which they are derived. Preferably said homologues have at least about 90 % sequence identity. To produce such homologues, amino acids present in the said protein can be replaced by other amino acids having similar properties, for example hydrophobicity, hydrophilicity, antigenicity, propensity to form or break α-helical structures or β-sheet structures, and so on. Conservative subsitution tables are well known in the art (see for example Creighton (1984) Proteins. W.H. Freeman and Company). An overview of physical and chemical properties of amino acids is given in Table 3.

Table 3. Properties of naturally occurring amino acids.

Charge properties/ hydrophobicity	**Side Group**	**Amino Acid**
nonpolar hydrophobic	aliphatic aliphatic, S-containing aromatic imino	ala, ile, leu, val met phe, trp pro
polar uncharged	aliphatic amide aromatic hydroxyl sulfhydryl	gly asn, gln try ser, thr cys
positively charged	basic	arg, his, lys
negatively charged	acidic	asp, gly

Two special forms of homology, orthologous and paralogous, are evolutionary concepts used to describe ancestral relationships of genes. The term "paralogous" relates to gene-duplications within the genome of a species leading to paralogous genes. The term "orthologous" relates to homologous genes in different organisms due to ancestral relationship. The present invention thus also relates to homologues, paralogues and orthologues of the proteins according to the invention.

Substitutional variants of a protein of the invention are those in which at least one residue in said protein amino acid sequence has been removed and a different residue inserted in its place. Amino acid substitutions are typically of single residues, but may be clustered depending upon functional constraints placed upon the polypeptide; insertions will usually be of the order of about 1-10 amino acid residues, and deletions will range from about 1-20 residues. Preferably, amino acid substitutions will comprise conservative amino acid substitutions, such as those described *supra.* Insertional amino acid sequence variants of a protein of the invention are those in which one or more amino acid residues are introduced into a predetermined site in said protein. Insertions can comprise amino-terminal and/or carboxy-terminal fusions as well as intra-sequence insertions of single or multiple amino acids. Generally, insertions within the amino acid sequence will be smaller than amino- or carboxy-terminal fusions, of the order of about 1 to 10 residues. Examples of amino- or carboxy-terminal fusion proteins or peptides include the binding domain or activation domain of a transcriptional activator as used in the yeast two-hybrid system, phage coat proteins, (histidine)$_6$-tag, glutathione S-transferase-tag, protein A, maltose-binding protein, dihydrofolate reductase, Tag•100 epitope, c-myc epitope, FLAG®-epitope, lacZ, CMP (calmodulin-binding peptide), HA epitope, protein C epitope and VSV epitope.

Deletion variants of a protein of the invention are characterized by the removal of one or more amino acids from said protein. Amino acid variants of a protein of the invention may readily be made using peptide synthetic techniques well known in the art, such as solid phase peptide synthesis and the like, or by recombinant DNA manipulations. The manipulation of DNA sequences to produce substitution, insertion or deletion variants of a protein are well known in the art. For example, techniques for making substitution mutations at predetermined sites in DNA are well known to those skilled in the art and include M13 mutagenesis, T7-Gen *in vitro* mutagenesis (USB, Cleveland, OH), QuickChange Site Directed mutagenesis (Stratagene, San Diego, CA), PCR-mediated site-directed mutagenesis or other site-directed mutagenesis protocols.

'Derivatives' of a protein of the invention are those peptides, oligopeptides, polypeptides, proteins and enzymes which may comprise additional naturally-occurring, altered glycosylated, acylated or non-naturally occurring amino acid residues compared to the amino acid sequence of a naturally-occurring form of said polypeptide. A derivative may also comprise one or more non-amino acid substitutents compared to the amino acid sequence of which it is derived, for example a reporter molecule or

other ligand, covalently or non-covalently bound to the amino acid sequence such as, for example, a reporter molecule which is bound to facilitate its detection.

The term 'cell cycle' means the cyclic biochemical and structural events associated with growth and with division of cells, and in particular with the regulation of the replication of DNA and mitosis. Cell cycle includes phases called: G0, Gap1 (G1), DNA synthesis (S), Gap2 (G2), and mitosis (M). Normally these four phases occur sequentially, however, the cell cycle also includes modified cycles wherein one or more phases are absent resulting in modified cell cycle such as endomitosis, acytokinesis, polyploidy, polyteny, and endoreduplication.

With 'recombinant DNA molecule' or 'chimeric gene' is meant a hybrid DNA produced by joining pieces of DNA from different sources through deliberate human manipulation.

The term 'expression' means the production of a protein or nucleotide sequence in the cell. However, said term also includes expression of the protein in a cell-free system. It includes transcription into an RNA product, post-transcriptional modification and/or translation to a protein product or polypeptide from a DNA encoding that product, as well as possible post-translational modifications. Depending on the specific constructs and conditions used, the protein may be recovered from the cells, from the culture medium or from both. For the person skilled in the art it is well known that it is not only possible to express a native protein but also to express the protein as fusion polypeptides or to add signal sequences directing the protein to specific compartments of the host cell, e.g., ensuring secretion of the peptide into the culture medium, etc. Furthermore, such a protein and fragments thereof can be chemically synthesized and/or modified according to standard methods described.

A 'vector' as used herein includes reference to a nucleic acid used for transfection or transformation of a host cell and into which a nucleic acid can be inserted. Expression vectors allow transcription and/or translation of a nucleic acid inserted therein. Expression vectors can for instance be cloning vectors, binary vectors or integrating vectors and typically contain control sequences as described *supra* to ensure expression in prokaryotic and/or eukaryotic cells. Advantageously, vectors of the invention comprise a selectable and/or scorable marker. Selectable marker genes useful for the selection of transformed plant cells, callus, plant tissue and plants are well known to those skilled in the art. For example, antimetabolite resistance provides the basis of selection for: the *dhfr* gene, which confers resistance to methotrexate (Reiss, Plant Physiol. (Life Sci. Adv.) 13 (1994), 143-149); the *npt* gene, which confers

resistance to the aminoglycosides neomycin, kanamycin and paromomycin (Herrera-Estrella, EMBO J. 2 (1983), 987-995); and *hpt*, which confers resistance to hygromycin (Marsh, Gene 32 (1984), 481-485). Additional selectable markers genes have been described, namely *trpB*, which allows cells to utilize indole in place of tryptophan; *hisD*, which allows cells to utilize histinol in place of histidine (Hartman, Proc. Natl. Acad. Sci. USA 85 (1988), 8047); *mannose-6-phosphate isomerase* which allows cells to utilize mannose (WO 94/20627) and *ornithine decarboxylase* which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine or DFMO (McConlogue, 1987, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed.) or *deaminase* from *Aspergillus terreus* which confers resistance to Blasticidin S (Tamura, Biosci. Biotechnol. Biochem. 59 (1995), 2336-2338). Useful scorable markers are also known to those skilled in the art and are commercially available. Advantageously, said marker is a gene encoding luciferase (Giacomin, Pl. Sci. 116 (1996), 59-72; Scikantha, J. Bact. 178 (1996), 121), green fluorescent protein (Gerdes, FEBS Lett. 389 (1996), 44-47) or ß-glucuronidase (Jefferson, EMBO J. 6 (1987), 3901-3907).

The vector or nucleic acid molecule according to the invention may either be integrated into the genome of the host cell or it may be maintained in some form extrachromosomally. In this respect, it is also to be understood that the nucleic acid molecule of the invention can be used to restore or create a mutant gene via homologous recombination or via other molecular mechanisms such as for example RNA interference (Paszkowski (ed.), Homologous Recombination and Gene Silencing in Plants. Kluwer Academic Publishers (1994)).

As used herein, a 'host cell' is a cell which contains a vector and supports the expression and/or replication of this vector. Host cells may be prokaryotic cells such as *E. coli* and *A. tumefaciens*, or it may be eukaryotic cells such as yeast, insect, amphibian, plant or mammalian cells. Preferably, host cells are monocotyledonous or dicotyledonous plant cells.

The term 'fragment of a sequence' or 'part of a sequence' means a truncated sequence of the original sequence referred to. The truncated sequence (nucleic acid or protein sequence) can vary widely in length; the minimum size being a sequence of sufficient size to provide a sequence with at least a comparable function and/or enzymatic activity of the original sequence referred to, while the maximum size is not critical. In some applications, the maximum size usually is not substantially greater than that required to provide the desired activity and/or function(s) of the original sequence.

Typically, the truncated amino acid sequence will range from about 5 to about 60 amino acids in length. More typically, however, the sequence will be a maximum of about 50 amino acids in length, preferably a maximum of about 30 amino acids. It is usually desirable to select sequences of at least about 10, 12 or 15 amino acids, up to about 20 or 25 amino acids.

Methods for alignment of nucleic acid and protein sequences were used herein to infer structural and functional similarities between aligned sequences. Methods for pairwise alignment of nucleic acid or protein sequences for comparative studies are well-known in the art. Algorithms have been described for optimal global sequence alignment, i.e. the alignment of two sequences over their entire length, (Smith and Waterman, Adv. Appl. Math. 2: 482 (1981)); and for finding local sequence similarities (Needleman and Wunsch, J. Mol. Biol. 48: 443 (1970); Pearson and Lipman, Proc. Natl. Acad. Sci. 85: 2444 (1988)). Examples of computerized implementations of such algorithms are: GAP (included in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wisconsin, USA) and FASTA (Lipman & Pearson, 1985). Multiple sequence alignment algorithms e.g. ClustalW (Higgins and Sharp, Gene 73:237-244 (1988)); PILEUP (Wisconsin Genetics Software Package) are based on a series of progressive, pairwise alignments between sequences and clusters of already aligned sequences to generate a final alignment.

The BLAST (Basic Local Alignment Search Tool) family of programs available at the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/BLAST/) was used to identify homologous sequences. As used herein, 'query' is a defined sequence that is used as a basis for alignment in for example, BLAST searches. A query may be a subset or the entirety of a specified sequence; for example it may be a full-length cDNA or a part thereof, a complete ORF or a part thereof. The BLAST software package includes: blastn to compare a nucleotide query sequence against a nucleotide sequence database; blastp to compare an amino acid query sequence against a protein sequence database; blastx to compare a nucleotide query sequence translated in all reading frames against a protein sequence database; tblastn to compare a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames; tblastx to compare the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database. Instead of identifying optimal global alignments, BLAST aims to identify regions of optimal local alignment, i.e. the alignment of some portion of two nucleic acid or protein sequences, to detect relationships among

sequences which share only isolated regions of similarity (Altschul et al., 1990). The E-value is used to indicate the expectation value. The lower the E-value, the more significant the alignment. See the NCBI website for a description of the alignment scores and statistics. In the present invention, the BLAST 2.0 suite of programs using default parameters was used (Altschul et al., Nucleic Acids Res. 25: 3389-3402 (1997)). Blast searches were performed on a local server or remotely through the NCBI server against public databases.

As used herein, 'sequence identity' in the context of two polypeptide sequences includes reference to the residues in the two sequences which are in the same position when aligned for maximum correspondence. With respect to polypeptide sequence alignment, scoring matrices used by the algorithms account for the fact that aligned residues which are not identical may be conservative amino acid substitutions, if amino acid residues are substituted for other amino acid residues with similar physicochemical properties. Sequences which differ by such conservative substitutions are said to have 'sequence similarity' and the percent identity may be adjusted upwards to correct for the conservative nature of the substitution. As used herein 'percentage of sequence identity' means the percentage calculated by determining the number of positions at which an identical amino acid residue occurs in both sequences (i.e. the number of matched positions), divided by the total number of residues in the smallest sequence, and multiplied by 100. AtCDKCIP1 homologous sequences were also identified using the complete AtCDKCIP1 protein sequence as query in a search against the Swissprot database using the Smith-Waterman alignment algorithm available at http://www.dna.affrc.go.jp/htbin/swp.pl. PHD domains in the AtCDKCIP1 protein were identified using the Pfam program available at http://www.sanger.ac.uk/cgi-bin/Pfam/nph-search.cgi (see http://www.sanger.ac.uk/Software/Pfam/help/scores.shtml for a discussion on the scores).

PEST regions in the AtCDKCIP1 protein were identified using the PESTfind program available at http://www.at.embnet.org/embnet/tools/bio/PESTfind/. The algorithm defines PEST sequences as hydrophilic stretches of amino acids greater than or equal to 12 residues in length. Such regions contain at least one P, one E or D and one S or T. They are flanked by lysine (K), arginine (R) or histidine (H) residues, but positively charged residues are not allowed within the PEST sequence (Rogers S., Wells R., Rechsteiner M.1986. Amino Acid Sequences Common to Rapidly Degraded Proteins: The PEST Hypothesis. Science 234, 364-368). PESTfind produces a score ranging

form about -50 to +50. By definition, a score above zero denotes a possible PEST region, but a value greater than +5 sparks real interest. Only PEST regions with values higher than 5 are described in the current application.

Nuclear localization signals were identified using the web-based Interpro service (http://www.ebi.ac.uk/interpro/scan.html).

AtCDKCIP1 homologous sequences were also identified using the complete AtCDKCIP1 protein sequence as query in a MPsrch_pp search (http://www.dna.affrc.go.jp/htdocs/MPsrch/MPsrch_pp.html) against the Swissprot database.

As used herein, the term 'plant' includes reference to whole plants, plant organs (such as leaves, roots, stems, etc.), seeds and plant cells and progeny of same. 'Plant cell', as used herein, includes suspension cultures, embryos, meristematic regions, callus tissue, leaves, seeds, roots, shoots, gametophytes, sporophytes, pollen, and microspores. The plants that can be used in the methods of the invention include all plants which belong to the superfamily *Viridiplantae*, in particular monocotyledonous and dicotyledonous plants including a fodder or forage legume, ornamental plant, food crop, tree, or shrub selected from the list comprising *Acacia spp., Acer spp., Actinidia spp.,Aesculus spp., Agathis australis, Albizia amara, Alsophila tricolor, Andropogon spp., Arachis spp, Areca catechu, Astelia fragrans, Astragalus cicer, Baikiaea plurijuga, Betula spp., Brassica spp., Bruguiera gymnorrhiza, Burkea africana, Butea frondosa, Cadaba farinosa, Calliandra spp, Camellia sinensis, Canna indica, Capsicum spp., Cassia spp., Centroema pubescens, Chaenomeles spp.,Cinnamomum cassia, Coffea arabica, Colophospermum mopane, Coronillia varia, Cotoneaster serotina, Crataegus spp., Cucumis spp., Cupressus spp., Cyathea dealbata, Cydonia oblonga, Cryptomeria japonica, Cymbopogon spp., Cynthea dealbata, Cydonia oblonga, Dalbergia monetaria, Davallia divaricata, Desmodium spp., Dicksonia squarosa, Diheteropogon amplectens, Dioclea spp, Dolichos spp., Dorycnium rectum, Echinochloa pyramidalis, Ehrartia spp., Eleusine coracana, Eragrestis spp., Erythrina spp., Eucalyptus spp., Euclea schimperi, Eulalia villosa, Fagopyrum spp., Feijoa sellowiana, Fragaria spp., Flemingia spp, Freycinetia banksii, Geranium thunbergii, Ginkgo biloba, Glycine javanica, Gliricidia spp, Gossypium hirsutum, Grevillea spp., Guibourtia coleosperma, Hedysarum spp., Hemarthia altissima, Heteropogon contortus, Hordeum vulgare, Hyparrhenia rufa, Hypericum erectum, Hyperthelia dissoluta, Indigo incarnata, Iris spp., Leptarrhena pyrolifolia, Lespediza spp., Lettuca spp., Leucaena leucocephala, Loudetia simplex, Lotonus bainesii, Lotus spp., Macrotyloma axillare, Malus spp.,*

Manihot esculenta, Medicago sativa, Metasequoia glyptostroboides, Musa sapientum, Nicotianum spp., Onobrychis spp., Ornithopus spp., Oryza spp., Peltophorum africanum, Pennisetum spp., Persea gratissima, Petunia spp., Phaseolus spp., Phoenix canariensis, Phormium cookianum, Photinia spp., Picea glauca, Pinus spp., Pisum sativum, Podocarpus totara, Pogonarthria fleckii, Pogonarthria squarrosa, Populus spp., Prosopis cineraria, Pseudotsuga menziesii, Pterolobium stellatum, Pyrus communis, Quercus spp., Rhaphiolepsis umbellata, Rhopalostylis sapida, Rhus natalensis, Ribes grossularia, Ribes spp., Robinia pseudoacacia, Rosa spp., Rubus spp., Salix spp., Schyzachyrium sanguineum, Sciadopitys verticillata, Sequoia sempervirens, Sequoiadendron giganteum, Sorghum bicolor, Spinacia spp., Sporobolus fimbriatus, Stiburus alopecuroides, Stylosanthos humilis, Tadehagi spp, Taxodium distichum, Themeda triandra, Trifolium spp., Triticum spp., Tsuga heterophylla, Vaccinium spp., Vicia spp.Vitis vinifera, Watsonia pyramidata, Zantedeschia aethiopica, Zea mays, amaranth, artichoke, asparagus, broccoli, brussel sprout, cabbage, canola, carrot, cauliflower, celery, collard greens, flax, kale, lentil, oilseed rape, okra, onion, potato, rice, soybean, straw, sugarbeet, sugar cane, sunflower, tomato, squash, and tea, amongst others. A particularly preferred plant is *Oryza sativa*.

The term 'transformation' as used herein, refers to the transfer of an exogenous polynucleotide into a host cell, irrespective of the method used for the transfer. The polynucleotide may be transiently or stably introduced into the host cell and may be maintained non-integrated, for example, as a plasmid, or alternatively, may be integrated into the host genome. The resulting transformed plant cell can then be used to regenerate a transformed plant in a manner known by a skilled person. *Agrobacterium*-mediated transformation or agrolistic transformation of plants, yeast, moulds or filamentous fungi is based on the transfer of part of the transformation vector sequences, called the T-DNA, to the nucleus and on integration of said T-DNA in the genome of said eukaryote. With "*Agrobacterium*" is meant a member of the *Agrobacteriaceae*, more preferably *Agrobacterium* or *Rhizobacterium* and most preferably *Agrobacterium tumefaciens*. With 'T-DNA', or transferred DNA, is meant that part of the transformation vector flanked by T-DNA borders which is, after activation of the *Agrobacterium vir* genes, nicked at the T-DNA borders and is transferred as a single stranded DNA to the nucleus of an eukaryotic cell. When used herein, with "T-DNA borders", 'T-DNA border region', or "border region" are meant either right T-DNA border (RB) or left T-DNA border (LB). Such a border comprises a core sequence flanked by a border inner region as part of the T-DNA flanking the border and/or a

border outer region as part of the vector backbone flanking the border. The core sequences comprise 22 bp in case of octopine-type vectors and 25 bp in case of nopaline-type vectors. One element enhancing T-DNA transfer has been characterised and resides in the right border outer region and is called *overdrive* (Peralta, Hellmiss et al., 1986;van Haaren, Sedee et al., 1987).

With 'T-DNA transformation vector' or 'T-DNA vector' is meant any vector encompassing a T-DNA sequence flanked by a right and left T-DNA border consisting of at least the right and left border core sequences, respectively, and used for transformation of any eukaryotic cell.

As used herein, 'transgenic plant' includes reference to a plant which comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a vector.

As used herein, the term 'heterologous' in reference to a nucleic acid is a nucleic acid that is either derived from a cell or organism with a different genomic background, or, if from the same genomic background, is substantially modified from its native form in composition and/or genomic environment through deliberate human manipulation. Accordingly, a heterologous protein although originating from the same species may be substantially modified by human manipulation.

'Transgenic' is used herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of the heterologous nucleic acid including those transgenics initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic.

The invention, now being generally described, will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention and are not intended to limit the invention.

All of the references mentioned herein are incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 - Sequence alignment of some CDK-like proteins related to animal CDK9. Arath;CDKC;1: CDKC;1 kinase from *Arabidopsis thaliana*; Arath;CDKC;2: CDKC;2 kinase from *Arabidopsis thaliana*; Medsa;CDKC;1: CDK protein from alfalfa (*Medicago sativa*); CDK9Hs: CDK9 protein from human (*Homo sapiens*); CDK9Dm: CDK9 protein from fruit fly (*Drosophila melanogaster*); CDK9Ce: CDK9 protein from *Caenorhabditis elegans*. The alignment was restricted to the region of the proteins that presented shared homology, for this reason the terminal ends have been omitted. Amino-acid residues identical in the six aligned proteins are indicated with asterisks, and the characteristic PITAL/IRE motif is boxed. The shadow regions in the CDKC;1 and CDKC;2 proteins correspond to the amino-acid residues which are not shared by both sequences.

Figure 2 - Sequence alignement of the cyclin T1 protein from Arabidopsis (CycT1At), mouse (CycT1Mou), human (CycT1Hs) and fruit fly (CycT1Dm). Amino-acid residues identical in all four protein sequences are highlighted by the asterisks. The alignment was restricted to the region of the proteins that presented sequence homology, for this reason the terminal ends have been omitted.

Figure 3 - Yeast two-hybrid interaction of Arabidopsis CDK proteins (CDKA;1, CDKB1;1 and CDKC;2) with cyclin T1 Arabidopsis homologue (CYCT). Yeast HF7c transformants were streaked on plates with (His^{+}) and without (His^{-}) histidine. Reconstitution of the GAL4 activity in the positive transformants restores the ability of the yeast to grow in histidine-lacking medium. Thus, showing that the plant cyclin T1 homologue protein is able to interact with Arath;CDKC;2 but not with Arath;CDKA;1 or Arath;CDKB1;1. 'cont' is the negative control, i.e. the empty bait vector pGBT9.

Figure 4 - *Arath;CDKC;*2 mRNA accumulation pattern in Arabidopsis flowers (4A through D) and radish roots (4E and F), as shown by *in situ* hybridization. In flowers *CDKC;*2 is confined to epidermic cells. *CDKC;*2 is developmentally regulated in flower tissues: at young stages transcripts are only visible in sepals (mainly the distal part) (Figure 4A and 4B), whereas in fully mature flowers the transcripts accumulate preferentially in petals and the expression in sepals slowly disappears (Figure 4C). In fully mature flowers *CDKC;*2 transcripts are also visible in the epidermis of the anthers

and the anther filament but never in the carpels (Figure 4C and 4D). *Arath;CDKC;2* transcripts were also observed in the endodermis of radish roots (Figure 4E and 4F).

Figure 5 - Sequence information on CDKC;2 and CDKC;2 interacting proteins and genes.

EXAMPLES

Example 1. Isolation of the Arath;*CDKC;2* gene

An expressed sequence tag (EST) encoding a CDK-like protein was initially identified by screening public databases (Burssens, Van Montagu et al., 1998). The full-length cDNA for this EST was subsequently cloned from an Arabidopsis cell suspension culture by 5'end amplification using the 5'end Capfinder kit (Clontech, Palo Alto, CA, USA). The full-length cDNA, designated Arath;*CDKC;2*, is 1738 bp long (SEQ ID NO: 1) and encodes a CDK-like protein of 505 amino acids (SEQ ID NO:2) with a calculated molecular weight of 56.7 Kd.

BLAST searches using SEQ ID NO: 1 as query against public genomic databases showed that the open reading frame of the Arath;*CDKC;2* cDNA is identical to the open reading frame of the predicted gene F18D22_40 located on BAC clone F18D22 (Acc. AL360334). The predicted protein of F18D22_40 (EMBL Acc. CAB96683.1 and PIR Acc. T150815) is annotated as a cdc2-like protein kinase.

The Arath;*CDKC;2* protein is highly homologous to three other CDK-like proteins in plants, all of which have the PITAIRE signature motif in the cyclin binding domain (Joubes, Chevalier et al., 2000)(see Figure 1 for Arath;*CDKC;2* and Medsa;*CDKC;2*):

(i) An *Arabidopsis thaliana* cDNA (GB Acc. AF360134) encoding a protein annotated as a cdc2-like protein kinase and renamed Arath;*CDKC*;1 (Joubes, Chevalier et al., 2000). The Arath;CDKC;1 protein has 92% amino acid sequence identity with the Arath;CDKC;2 protein.

(ii) A *Medicago sativa* cDNA (EMBL Acc. X97314) encoding a protein annotated as a cdc2 kinase homologue, and renamed Medsa;*CDKC;2* (Joubes, Chevalier et al., 2000). The protein encoded by Medsa; *CDKC;2* has about 80% peptide sequence identity with the Arath;CDKC;2 protein.

(iii) A partial protein from *Pisum sativum* (Acc. CAA39904) and renamed Pissa;CDKC;1 (Joubes, Chevalier et al., 2000).

BLAST searches also revealed sequence similarity between the Arath;*CDKC;2* protein and animal CDK9. This is illustrated in Figure 1 which shows a partial protein alignment of CDK9 from human (CDK9Hs), *Drosophila* (CDK9Dm), *Caenorhabditis* (CDK9Ce) and the Arath;CDKC;2 and Medsa;CDKC;2 protein. The Arath;*CDKC;2* protein has

50% sequence identity with CDK9 from human and, among all plant proteins, is the most closely related to human CDK9. The Arath;*CDKC;2* protein has a potential bipartite nuclear localization signal at position 350-367 as identified herein in a PROSITE Profile search (http://www.isrec.isb-sib.ch/software/PFSCAN_form.html), suggesting that this kinase accumulates and has a function inside the nucleus. Human CDK9 is part of the positive transcription elongation factor P-TEFb (Marshall et al., 1996; Price 2000).

Example 2. The CYCT1At cDNA was isolated in a two-hybrid screen using the Arath;CDKC;2 protein as bait

To identify the cyclin regulator of Arath;CDKC;2 and other protein interactors of Arath;CDKC;2, a yeast two-hybrid screen was performed using the Arath;CDKC;2 as bait. The bait construct was prepared by cloning a PCR amplified Arath;*CDKC;2* fragment cut with EcoRI/BamHI into the EcoRI/ BamHI sites of the yeast two-hybrid bait vector pGBT9. The two-hybrid prey library was derived from *Arabidopsis thaliana* (De Veylder, Segers et al., 1997a). Vectors and strains were from the Matchmaker two-hybrid system kit (Clontech, Palo Alto, CA, USA). Two-hybrid assays and screens were performed according to the manufacturer's protocol. Positive clones were identified by growth on histidine lacking medium. Prey plasmids from positive clones were isolated and sequenced as previously described (De Veylder, Segers et al, 1997a). These cDNA sequences were subsequently used in BLAST searches against public databases.

In this way, a cDNA was isolated encoding a protein that showed high sequence homology to the cyclin T from mouse (Acc. AAD17205). This cDNA was designated CYCT1At for cyclin T1 of *Arabidopsis thaliana.* The full-length cDNA and peptide sequence is represented as SEQ ID NO:3 and SEQ ID NO:4 respectively. The sequence alignment of Figure 2 illustrates the sequence similarity between CYCT1At and cyclin T from human, mouse and *Drosophila*.

The identification of a cyclin T-like protein as the cyclin regulator of Arath;CDKC;2, as disclosed herein, may indicate that the Arath;CDKC;2/CYCT1At heterodimer is structurally and functionally homologous to the human CDK9/cyclinT pair, which is involved in transcription regulation.

BLASTP searches using the complete ORF of CYCT1At as query against the protein sequence database identified a nearly identical protein (GB Acc. AAD 46000.1) that

differed in only one amino acid position from CYCT1At (P at position 277 substituted by L). When using the *CYCT1At* nucleotide sequence as query against the nucleotide sequence database in BLASTN searches, a coding and genomic sequence was identified that is identical to the *CYCT1At* sequence except in one position: the coding sequence (Acc. AF344323) has a C at position 830 which is T in CYCT1At. The coding sequence is derived from the predicted gene T17H3.12 and the encoded protein is annotated in the public database as an unknown protein that contains similarity to the silencing mediator of retinoic acid and thyroid hormone receptor alpha and cyclin T1 from *Mus musculus*.

Example 3. The Arabidopsis Arath;CDKC;2 and CYCT1At proteins specifically interact with each other in a yeast two-hybrid assay

Cyclin-dependent kinases form a conserved family of protein kinases in eukaryotes. Based on structural and functional properties, five classes of CDKs have been recognized in plants: *CDKA*, *CDKB*, *CDKC*, *CDKD*, and *CDKE*. CDKs require a functional association with a cyclin partner to be active. To a large extent it is the cyclin partner that defines the substrate specificity of the complex. Therefore, formation of a specific CDK/cyclin pair can yield information about its functionality. The *CDKA* and *CDKB* class comprises genes that are involved in cell cycle regulation. No functional information is available for plant *CDKC* genes.

To address the specificity of the interaction of CYCT1At with the Arath;*CDKC;2* protein, two-hybrid assays were performed with Arath;*CDKC;2* and with a member of the *CDKA* and *CDKB* class. Two-hybrid bait vectors containing the Arath;CDKA;1, Arath;CDKB;1 or Arath;CDKC;2 were constructed as described (De Veylder, Segers et al., 1997b). The CYCT1At prey was constructed by inserting the coding region (position 1 to 954 in SEQ ID NO: 3) into a gateway vector (GATEWAY Cloning Technology; Life Technologies), containing the GAL4 activation domain. Insertion of the CYCT1At fragment was done by recombination between the *att*B sequence of the gateway vector and the CYCT1At fragment, which was amplified by PCR using primers containing terminal *att*B sites (according to GATEWAY Cloning Technology protocol book). Plasmids encoding bait and prey fusion proteins were co-transformed into the yeast reporter strain HF7c and interactions between the two proteins were assayed by the ability of the co-transformed strain to grow on histidine lacking medium.

As shown in Figure 3, the CYCT1At protein interacts with Arath;CDKC;2 but not with Arath;CDKA;1 or Arath;CDKB;1 as demonstrated by growth on histidine lacking medium only for the combination CYCT1At and Arath;CDKC;2. This demonstrates that the CYCT1At protein specifically interacts with Arath;CDKC;2 but not with a member of class A and B CDKs.

Example 4. The Arath;CDKC;2 protein also interacts with proteins involved in transcription, RNA processing, plastid development and photosynthesis.

In addition to CYCT1At, five other clones were isolated as interactors of the Arath;CDKC;2 protein in the two-hybrid screen described in Example 2. Prey plasmids were isolated from these positive interactors and the cDNA inserts were partially sequenced. Translated cDNA sequences were used in BLASTP searches to identify the encoded proteins.

The proteins that were identified as interactors of Arath;CDKC;2 include proteins that play a role in photosynthesis and chloroplast development as well as proteins involved in transcription processes. A description of the isolated cDNAs with results of the BLAST searches is described below:

1. Ribulose-bisphosphate carboxylase/oxygenase activase

The cDNA insert of a second Arath;CDKC;2 interacting prey plasmid was partially sequenced and this sequence is represented as SEQ ID NO:5. This sequence is 524 bp long, has a startcodon at position 98 and encodes a partial protein of 142 amino acids represented as SEQ ID NO:6.

BLASTP searches using SEQ ID NO:6 as query against the protein database identified this protein as a ribulose-bisphosphate carboxylase/oxygenase (rubisco) activase-like protein. The specifications for the first retrieved alignment between query and subject (Acc.TO1003) are as following: Expect = 3e-38; Identities = 93/143 (65%), Positives = 104/143 (72%), Gaps = 2/143 (1%).

Rubisco activase is a regulator of rubisco which itself is involved in the fixation of atmospheric CO_2. Rubisco activase controls the overall process of photosynthesis by making rubisco activity responsive to light intensity (Jensen, 2000).

2. DAG-like protein

The cDNA insert of a third Arath;CDKC;2 interacting prey plasmid was partially sequenced and this sequence is represented as SEQ ID NO:7. This sequence is 657 bp long and encodes a partial protein of 219 amino acids represented as SEQ ID NO:8. BLASTP searches using SEQ ID NO:8 as query against the protein database indicated that this peptide sequence is 100% identical to an internal part of a protein of *Arabidopsis thaliana* (GB Acc BAA97063.1) that is annotated as containing similarity to DAG protein. The sequence of this protein is represented as SEQ ID NO:9. The peptide sequence of SEQ ID NO:8 is identical to the protein sequence represented as SEQ ID NO:9 from position 24 to position 239 (note that the first three AA of SEQ ID No.8 are translated vector sequence). Therefore, this Arath;CDKC;2 interactor was identified as a DAG-like protein. The DAG (differentiation and greening) protein was originally identified in *Antirrhinum majus* by transposon tagging and the gene is required for chloroplast differentiation and palissade development (Chatterjee, Sparvoli et al., 1996). Expression of DAG is essential for expression of plastid and nuclear genes affecting the chloroplasts such as rubisco activase and also for expression of the plastidial gene encoding the beta subunit of plastidial RNA polymerase.

3. ribonucleoprotein

The cDNA insert of a fourth Arath;CDKC;2 interacting prey plasmid was partially sequenced and this sequence is represented as SEQ ID NO:10. This sequence is 639 bp long, does not have a startcodon, neither a stopcodon and encodes a partial protein of 213 amino acids represented as SEQ ID NO:11.

BLASTP searches using SEQ ID NO:11 as query against the protein database identified this protein as a probable ribonucleoprotein (RNP) with following specifications for the first retrieved alignment between query and subject (Acc.G71404): Expect = e-126, identities = 210/213 (98%), positives = 211/213 (98%). Therefore, this Arath;CDKC;2 interactor was identified as a ribonucleoprotein. Recent data showed a functional coupling between RNA polymerase II transcription and RNA processing by RNP proteins (Bentley, 1999). The finding that Arath;CDKC;2 interacts with an RNP confirms that the Arath;CDKC;2 protein and/or the Arath;CDKC;2 /CYCT1At protein complex is implicated in transcription regulation.

4. AtCDKCIP1

The cDNA insert of a fifth Arath;CDKC;2 interacting prey plasmid was partially sequenced and this sequence is represented as SEQ ID NO:12. This sequence is 589 bp long including a poly(A) tail of 22 nucleotides, has a stopcodon located at position 379, and encodes a polypeptide of 126 amino acids represented as SEQ ID NO:13.

BLASTP searches using SEQ ID NO:13 as query against the protein database revealed that this peptide sequence was identical to the carboxy-terminal part of the protein encoded by the predicted gene MTE17.10 (Acc. AB015479) located on chromosome V of *Arabidopsis thaliana*. This protein is annotated as an unknown protein (Acc. BAB08556) in the database. In addition, the 3'untranslated region of SEQ ID NO:12 was 100% identical to the 3'UTR of gene MTE17.10 (data not shown). Therefore, the gene product of MTE17.10 is an interactor of the Arath;CDKC;2, as disclosed herein and is designated AtCDKCIP1 for *Arabidopsis thaliana* CDKC Interacting Protein 1. The peptide sequence of AtCDKCIP1 is represented as SEQ ID NO:14. The AtCDKCIP1 protein is 1332 amino acids long.

As disclosed herein, AtCDKCIP1 comprises five potential PEST sequences as determined by PESTfind. Three highly significant PEST regions, i.e. with a value greater than 5, are located at position 0-28 (MTFVDDDEEEDFSVPQSASNYYFEDDDK SEQ ID NO 18; Pest-find score 7.43); at position 589-604 (KEPGSEIPTLDNDSQR SEQ ID NO 19; Pest find score 8.26) and at position 1293-1310 (HDFPLPPPPPSDFEMSPR SEQ ID NO 20; Pest find score 8.28). PEST regions serve as proteolytic signals, indicating that AtCDKCIP1 is subject to specific protein degradation mechanisms.

AtCDKCIP1 further contains putative bipartite nuclear localization signals (at position 493-510 and 611-628), as identified in an InterPro search (http://www.ebi.ac.uk/interpro/scan.html) using the complete AtCDKCIP1 peptide sequence as query. The AtCDKCIP1 protein therefore accumulates in the nucleus and/or has a function in the nucleus.

The AtCDKCIP1 protein also has two potential PHD domains. The first PHD domain starts at position 224 and ends at position 281 (e-value 0.005). The second PHD domain starts at position 284 and ends at position 350 (e-value 0.002). The PHD finger is a C4HC3 zinc finger-like motif found in nuclear proteins thought to be involved in chromatin-mediated transcriptional regulation. The PHD finger motif is reminiscent of, but distinct from, the C3HC4 type RING finger. The function of this domain is not yet known but in analogy with the LIM domain it could be involved in protein-protein

interaction and be important for the assembly or activity of multi-component complexes involved in transcriptional activation or repression. In similarity to the RING finger and the LIM domain, the PHD finger is thought to bind two zinc ions.

The AtCDKCIP1 furthermore shares significant homology with DNA binding proteins identified in a MPsrch_pp search using SEQ ID NO:14 as query against the Swissprot database. The first 4 retrieved alignments are listed below; three of the identified proteins are DNA binding proteins and a fourth protein is a transcription factor.

RESULT 1

ID CHD3_HUMAN STANDARD; PRT; 1944 AA.
DE CHROMODOMAIN HELICASE-DNA-BINDING PROTEIN 3 (CHD-3) (MI-2 AUTOANTIGEN
DE 240 KDA PROTEIN) (MI2-ALPHA).

DB 1; Score 164; Match 40.4%; QryMatch 1.4%; Pred. No. 5.39e-14;
Matches 19;Conservative 13;Mismatches 12;Indels 3;Gaps 3;

```
            .*. *       *.*  .  .*.**    *    *   .*.    ***.***  .  .*.  *.*
Db  435 EEEEYEEE G EEEGEKEEEDDHMEY-CRVCKDGGELLCCD-ACI SSYH 479
Qy  201 DEDTYVASDEDELD- DEDDDFFESV CA ICDNGGEI LCCEGSCLRSFH 246
```

RESULT 2

ID CHD4_HUMAN STANDARD; PRT; 1912 AA.
DE CHROMODOMAIN HELICASE-DNA-BINDING PROTEIN 4 (CHD-4) (MI-2 AUTOANTIGEN
DE 218 KDA PROTEIN) (MI2-BETA).

DB 1; Score 147; Match 45.7%; QryMatch 1.3%; Pred. No. 3.56e-10;
Matches 16;Conservative 9;Mismatches 8;Indels 2;Gaps 2;

```
            .*. .***      *    *  .*   .***.***.   .*   *.*
Db  440 DLEEEDDHHMEF-  CRVCKDGGELLCCD- TCPSSYH 472
Qy  212 ELDDEDDD F FESVCAI CDNGGE I LCCEGSCLRSFH 246
```

RESULT 3

ID CHDM_DROME STANDARD; PRT; 1982 AA.
DE CHROMODOMAIN HELICASE-DNA-BINDING PROTEIN MI-2 HOMOLOG (DMI-2).

DB 1; Score 145; Match 40.0%; QryMatch 1.2%; Pred. No. 9.68e-10;
Matches 16;Conservative 11;Mismatches 12;Indels 1;Gaps 1;

```
            .*         . .****    .    *   .*   .***.* **.   **    . .*.
Db  423 ADGGAAEEEDDDEHQEFCRVCKDGGELLCCD-SCPSAYHT 461
Qy  208 S DEDELDDEDDDFF ESVCA I CDNGGEILCCEGSCLRSFHA 247
```

RESULT 4

ID TF1G_HUMAN STANDARD; PRT; 1127 AA.
DE TRANSCRIPTION INTERMEDIARY FACTOR 1-GAMMA (TIF1-GAMMA) (RFG7 PROTEIN).

DB 1; Score 140; Match 51.4%; QryMatch 1.2%; Pred. No. 1.14e-08;
Matches 18; Conservative 7;Mismatches 9;Indels 1;Gaps 1;

```
          . . * * *   *   * * . * . * * * . . * * **   * .   **  *
Db   879 NNKDDDPNEDWCAVCQNGGDLLCCE-KCPKVFHLT 912
Qy   214 DDEDDD FFE SV CA ICDNGGE ILCCEGSCLRSFHAT 248
```

Collectively, the data disclosed in this invention indicate that the AtCDKCIP1 interactor of the Arath;CDKC;2 is a nuclear protein involved in transcription regulation processes. This finding provides further evidence that Arath;CDKC;2 and/or a multiprotein complex containing Arath;CDKC;2 is implicated in transcription regulation processes.

5. AtGT-1

The cDNA insert of a sixth Arath;CDKC;2 interacting prey plasmid was partially sequenced and this sequence is represented as SEQ ID NO:15. This sequence is 664 bp, has a startcodon located at position 24-26, and encodes a polypeptide of 213 amino acids represented in SEQ ID NO:16.

BLASTP searches using SEQ ID NO:16 as query against the protein database revealed that this peptide is identical to the carboxy-terminal part of the GT-1 protein from *Arabidopsis thaliana* encoded by the *AtGT-1* gene. The sequence of the protein encoded by *AtGT-1* is represented as SEQ ID No 17. The AtGT-1 protein is a DNA binding protein and a regulator of light-activated expression of the gene encoding the small subunit of ribulose bisphosphate carboxylase (Hiratsuka, Wu et al., 1994; Zhou, 1999). The interaction of Arath;CDKC;2 with the transcription factor AtGT-1 therefore indicates that Arath;CDKC;2 and/or a protein complex containing Arath;CDKC;2 may be involved in light-regulated transcription processes.

Example 5. Expression analysis of the *Arath;CDKC;1* and *Arath;CDKC;2* and *CYCT1At* gene in *Arabidopsis thaliana* tissues

The expression of *Arath;CDKC;1*, *Arath;CDKC*;2 and *CYCT1At* was examined by real-time PCR. Total RNA was extracted from young seedlings, roots, rosettes, stems and flowers of *Arabidopsis thaliana* (L.) Heynh. ecotype Columbia according to standard protocols. Two microgram of each sample were reverse-transcribed into cDNA using the SuperScript First-Strand Synthesis System for RT-PCR (Gibco BRL; Life Technologies). Semi-quantitative RT-PCR amplification of the cDNA was carried out in

a LightCycler real-time PCR (Roche Diagnostics), using gene-specific primers: for *Arath;CDKC;1*: 5'-ACATTCTCGTTTACCTCCACAG-3' (SEQ ID NO 21) as forward and 5'-AAAATCACAACTGCCTTAAAGAC-3' (SEQ ID NO 22) as reverse primer; for *Arath;CDKC;2*: 5'-ACCCAGCCACAACTTCTATG-3' (SEQ ID NO 23) as forward and 5'-CTAGTATCACATTAAATGTAAGAGTAAG-3' (SEQ ID NO 24) as reverse primer; for *CYCT1At* 5'-TGTCGTTGTAGCGTCTTATG-3' (SEQ ID NO 25) as forward and 5'-TCCTTCTGTCCACTTCTATC-3' (SEQ ID NO 26) as reverse primer. The amount of target cDNA used for PCR was standardized by quantification of actin 2 transcripts present in all the samples. Independent experiments showed a maximum of 20% error. The results are summarized in Table 1 and showed that *Arath;CDKC;1*, *Arath;CDKC*;2 and *CYCT1At* transcripts, although present in all tested organs, were most abundant in flower tissues. The amount of transcripts detected in flowers for the three genes was about two-fold higher than in all other tested organs.

Table 1. Semi-quantitative transcript analyses by real-time RT-PCR.

	Arath;CDKC;1	***Arath;CDKC;2***	***CYCT1At***
seedlings	141,30	12,160	17,430
root	97,340	13,330	16,290
rosettes	103,70	11,860	25,090
stems	130,50	09,130	30,620
flowers	240,60	23,480	50,140

Expression of *Arath;CDKC;2* was also analyzed by Northern analysis. Hybridization with an antisense riboprobe revealed the existence of two similar sized transcripts of approximately 1.8Kb, which correspond to the *Arath;CDKC;1* and *Arath;CDKC*;2 transcripts (data not shown).

Example 6. In situ hybridization of *Arath;CDKC;2* reveals a tissue-specific expression pattern that is restricted to terminally differentiated tissues.

The expression pattern of the *Arath;CDKC;2* and *CYCT1At* gene was studied by *in situ* RNA hybridization of *Arabidopsis thaliana* tissues and radish roots. Plant material was fixed in 3% glutaraldehyde in 0.1 M cacodylate buffer, pH 7.2 (12h at 4°C). Fixed tissues were dehydrated through standard ethanol series, and embedded in paraffin. Tissue serial sections of 10 µm were attached to coated microscope slides. ^{35}S-UTP-labelled sense (control) and antisense RNA probes for *Arath;CDKC;2* and *CYCT1At* were generated by *in vitro* transcription with T7 and Sp6 RNA polymerases, according to the manufacturer's protocol (Boehringer-Mannheim; Germany). Full-length transcripts were reduced to 300 bp fragments through alkaline hydrolysis. Plant material was hybridized overnight at 42°C with the appropriated anti-sense and control probes ($5x10^6$ c.p.m. per slide). All hybridization procedures were performed as described (de Almeida Engler, de Groodt et al., 2001). Autoradiographs were taken under dark-field illumination in an optic microscope Diaplan (Leitz, Heerbrugg, Switzerland).

In flowers, the transcript was mainly confined to the epidermic cell layer in petals (both inner and outer epidermis) and sepals (only outer epidermis) (Figure 4A). Furthermore, the *Arath;CDKC;2* gene appears to be developmentally regulated in flowers since at young stages transcripts were only visible in sepals (mainly distal part) (Figure 4A and B), whereas in fully mature flowers the transcripts accumulated preferentially in petals and the expression in sepals slowly disappeared (Figure 4C). *Arath;CDKC;2* transcripts are also visible in the epidermis of the anthers and the anther filament, but only in fully mature flowers (Figure 4C and D). Conversely, *Arath;CDKC;2* mRNA was never detected in carpels (Figure 4C and 4D). However, *Arath;CDKC;2* transcripts were visible afterwards in the outer epidermis of siliques (data not shown). As shown in Figure 4E and 4F, expression of the *Arath;CDKC;2* gene in roots was confined to the endodermic cell layer. Importantly, *Arath;CDKC;2* gene expression was not observed in meristematic tissues. Also, no expression was detected in leaves at any developmental stage. These results demonstrated that the Arath;CDKC;2 protein is not directly involved in cell division control.

References

Bentley, D. Coupling RNA polymerase II transcription with pre-mRNA processing (1999). Curr Opin Cell Biol **11**, 347-351.

Burssens, S, Van Montagu, M., and Inze, D. The cell cycle in *Arabidopsis* (1998). Plant Physiol.Biochem. **36**, 9-19.

Chatterjee, M., Sparvoli, S., Edmunds, C., Garosi, P., Findlay, K., and Martin, C. (1996). DAG, a gene required for chloroplast differentiation and palisade development in *Antirrhinum majus. EMBO J* **15**, 4194-4207.

de Almeida Engler, J, de Groodt, R, Van Montagu, M, and Engler, G. (2001). *In situ* hybridization to mRMA of *Arabidopsis* tissue sections. Methods **23**, 325-334. De Veylder, L., Segers, G, Glab, N, Van Montagu, M., and Inze, D. (1997a). Identification of proteins interacting with the Arabidopsis Cdc2aAt protein. J EXP BOTANY **48**, 2113-2114.

De Veylder, L., Segers, G., Glab, N., Casteels, P., Van Montagu, M., and Inze, D. (1997b). The Arabidopsis Cks1At protein binds the cyclin-dependent kinases Cdc2aAt and Cdc2bAt. *FEBS Lett.* **412**, 446-452.

Hiratsuka, K, Wu, X., Fukuzawa, H, and Chua, NH. Molecular dissection of GT-1 from Arabidopsis. (1994). Plant Cell **6**, 1805-1813.

Jensen, RG (2000). Activation of Rubisco regulates photosynthesis at high temperature and CO_2. Proc. Natl. Acad. Sci. **97**, 12937-12938.

Joubes,J., Chevalier,C., Dudits,D., Heberle-Bors,E., Inze, D., Umeda, M., and Renaudi, J. P. (2000). CDK-related protein kinases in plants. *Plant Mol.Biol* **43**, 607-620.

Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl- terminal domain kinase. *J.Biol Chem.* **271**, 27176-27183.

Mironov, V., De Veylder, L., Van Montagu, M., and Inze, D. (1999). Cyclin-dependent kinases and cell division in plants- the nexus. *Plant Cell* **11**, 509-522.

Morgan. Cyclin-dependent kinases: engines, clocks, and microprocessors. (1997). Annu Rev Cell Dev Biol **13**, 261-291.

Price, D. H. (2000). P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol.Cell Biol* **20**, 2629-2634.

Zhou, D-X. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. (1999). Trends in Plant Science **4**, 210-214.

Claims

1. A method for altering growth and/or yield characteristics of a plant or plant cell comprising modulating the expression in a plant or plant cell of at least one first nucleic acid encoding a plant CDKC kinase, a homologue or a derivative thereof or an enzymatically active fragment thereof and/or at least one second nucleic acid encoding a CDKC kinase interacting protein, a homologue or a derivative thereof or an enzymatically active fragment thereof.

2. A method according to claim 1 said method comprising modulating transcription regulation.

3. A method according to claim 1 said method comprising modulating photosynthesis and/or chloroplast development.

4. A method according to claim 1 for enhancing photosynthetic capacity of a plant or a plant cell.

5. A method according to claim 1 for increasing the number of flowers and/or seeds and/or fruits of a plant.

6. A method according to any of claims 1 to 5 wherein said plant CDKC kinase is represented by SEQ ID NO 2 and wherein said CDKC kinase interacting protein is chosen from the polypeptides represented by any of SEQ ID NOs 4, 6, 8, 9, 11, 13, 14, 16 or 17.

7. A method according to any of claims 1 to 6 comprising stably integrating into the genome of said plant at least one of said first or second nucleic acids in an expressible form.

8. A method according to any of claims 1 to 7 comprising downregulation of expression of said first or second nucleic acid.

9. A method according to claim 8 comprising stably integrating into the genome of said plant at least one nucleic acid causing downregulation of expression of said first or second nucleic acid.

10. A method according to claim 9 wherein said nucleic acid comprises at least part of an antisense version of said first or said second nucleic acid as defined in claim 1.

11. A method according to claim 10 wherein said first nucleic acid is represented in SEQ ID NO 1 and said second nucleic acid is chosen from the group of nucleic acids represented in SEQ ID NOs 3, 5, 7, 10, 12 or 15.

12. The method according to any of claims 1 to 11 comprising downregulation of expression of a nucleic acid encoding CYCT1At represented by SEQ ID NO 4, or a homologue thereof.

13. The method according to any of claims 1 to 12 wherein said plant CDKC kinase is represented by SEQ ID NO 2 or a derivative thereof or an enzymatically active fragment thereof and wherein said CDKC kinase interacting protein is CYCT1At represented by SEQ ID NO 4 or a derivative thereof or an enzymatically active fragment thereof.

14. A method for the production of a transgenic plant having altered growth and/or yield characteristics comprising:

   - transforming a plant cell with a DNA construct comprising a (i) gene promoter sequence, (ii) at least one open reading frame encoding at least one functional portion of a CDKC kinase, or a homologue or a derivative thereof, and/or (iii) at least one second open reading frame encoding at least one functional portion of a CDKC kinase interacting protein, a homologue or a derivative thereof, to provide a transgenic cell;
   - providing means for altering the expression of said nucleic acid, and
   - cultivating the transgenic cell under conditions promoting regeneration and mature plant growth.

15. A method for the production of a transgenic plant having altered growth and/or yield characteristics comprising:

   - transforming a plant cell with a DNA construct comprising at least one nucleic acid as defined in any of claims 8 to 10 under the control of a promoter sequence, to provide a transgenic cell; and
   - cultivating the transgenic cell under conditions promoting regeneration and mature plant growth.

16. A method according to any of claims 1 to 15 wherein said plant or plant cell is derived from rice (*Oryza sativa*).

17. A transgenic plant obtainable by any of the methods of claims 1 to 16.

18. A method for identifying and obtaining compounds that interfere with the interaction between a CDKC kinase and a CDKC kinase interacting protein comprising the steps of :

(a) providing an expression system wherein a CDKC kinase, a homologue or a derivative thereof or a fragment thereof and a CDKC kinase interacting protein, a homologue or a derivative thereof or a fragment thereof are expressed,
(b) interacting at least one compound with the complex formed by the expressed polypeptides as defined in (a), and,
(c) measuring the effect of said compound on the binding between the interacting proteins as defined in (a) or measuring the activity of said complex;
(d) optionally identifying said compound.

19. The method of claim 18 wherein said compound inhibits the activity of said protein complex or inhibits the formation of a complex between said proteins.

20. The method of claim 18 wherein said compound enhances the activity of said protein complex or promotes the formation of a complex between said proteins or influences the activity of said complex.

21. A compound obtainable by any of the methods of claims 18 to 20.

22. Use of a compound identified by means of any of the methods of claims 18 to 20 as a plant growth regulator.

23. Use of a compound identified by means of any of the methods of claims 18 to 20 as a plant herbicide.

24. A method for production of a plant growth regulator or herbicide composition comprising the steps of the method of any of claims 18 to 20 and formulating the compounds obtained from said steps in a suitable form for the application in agriculture or plant cell or tissue culture.

25. A method for the design of or screening for growth-promoting chemicals or herbicides comprising the use of a nucleic acid encoding a protein as defined in claim 6.

26. Use of a nucleic acid encoding a protein as defined in claim 6 for increasing yield.

27. Use of a nucleic acid encoding a protein as defined in claim 6 for stimulating growth.

28. Use of a nucleic acid encoding a protein as defined in claim 6 for increasing the number of flowers and/or seeds and/or fruits per plant.

29. Use of a nucleic acid encoding a protein as defined in claim 6 for modulating transcription regulation processes.

30. Use of a nucleic acid encoding a protein as defined in claim 6 for enhancing photosynthetic capacity.

```
                       ** **   * *  *           ** *       ** *** **** *** *
Arath;CDKC;1   28  KLEQIGEGTY GQVYMAKEIK TGE.IVALKK IRMDNEREGF PITAIREIKI
Arath;CDKC;2   28  KLEQIGEGTY GQVYMAKEIK TGE.IVALKK IRMDNEREGF PITAIREIKI
Medsa;CDKC;1   27  KLEQIGEGTY GMVYMAREIE TGE.IVALKK IRMDNEREGF PITAIREIKI
     CDK9Dm    52  KVAKIGQGTF GEVFKAREKK GNKKFVAMKK VLMDNEKEGF PITALREIRI
     CDK9Hs    21  KLAKIGQGTF GEVFKARHRK TGQK.VALKK VLMENEKEGF PITALREIKI
     CDK9Ce     1  MLDQIGEGTY GQVYKAVNNL TGE.QVALKR VRLENEKEGF PITAIREIKI

                   *  * ***       *                                **     ***
Arath;CDKC;1   77  LKKLHHENVI HLKEIVTSPG RDRDDQGKPD NNKYKGGIYM VFEYMDHDLT
Arath;CDKC;2   77  LKKLHHENVI QLKEIVTSPG RDRDDQGKPD NNKYKGGIYM VFEYMDHDLT
Medsa;CDKC;1   76  LKKLHHENVI KLKEIVTSPG PEKDDQGRPD GNKYKGGIYM VFEYMDHDLT
     CDK9Dm   102  LQLLKHENVV NLIEICRTKA TAT....... .NGYRSTFYL VFDFCEHDLA
     CDK9Hs    70  LQLLKHENVV NLIEICRTKA SPY....... .NRCKGSIYL VFDFCEHDLA
     CDK9Ce    50  LRQLHHKNIV RLMDIVI... ...DDISMDE LKRTRANFYL VFEYVDHDLI

                   **         *     *      ** ** *  *     ****  *   * *
Arath;CDKC;1  127  GLADRPGL.R FTVPQIKCYM KQLLTGLHYC HVNQVLHRDI KGSNLLIDNE
Arath;CDKC;2  127  GLADRPGL.R FTVPQIKCYM KQLLTGLHYC HVNQVLHRDI KGSNLLIDNE
Medsa;CDKC;1  126  GLADRPGM.R FTVPQIKCYM RQLLTGLHYC HVNQVLHRDI KGSNLLIDNE
     CDK9Dm   144  GLLSNM.NVK FSLGEIKKVM QQLLNGLYYI HSNKILHRDM KAANVLITKH
     CDK9Hs   112  GLLSNV.LVK FTLSEIKRVM QMLLNGLYYI HRNKILHRDM KAANVLITRD
     CDK9Ce    94  GLLESKELVD FNKDQICSLF KQLLEGLAYI HNTGFLHRDI KCSNILVNNK

                   * ** ** ** **                  **** ********** **    ***
Arath;CDKC;1  176  GNLKLADFGL ARSYS...HD HTGNLTNRVI TLWYRPPELL LGATKYGPAI
Arath;CDKC;2  176  GNLKLADFGL ARSYS...HD HTGNLTNRVI TLWYRPPELL LGATKYGPAI
Medsa;CDKC;1  175  GNLKLADFGL ARSFS...NE HNANLTNRVI TLWYRPPELL LGTTRYGPAV
     CDK9Dm   193  GILKLADFGL ARAFSIPKNE SKNRYTNRVV TLWYRPPELL LGDRNYGPPV
     CDK9Hs   161  GVLKLADFGL ARAFSLAKNS QPNRYTNRVV TLWYRPPELL LGERDYGPPI
     CDK9Ce   144  GELKIADLGL ARLWE...KE .SRLYTNRVI TLWYRPPELL LGDERYGPAI

                   * *  **    *     *    *  *  **  *   **
Arath;CDKC;1  223  DMWSVGCIFA ELLNGKPILP GKTENEQLNK IYELCGSPDE SNWPGVSKMP
Arath;CDKC;2  223  DMWSVGCIFA ELLHAKPILP GKNEQEQLNK IFELCGSPDE KLWPGVSKMP
Medsa;CDKC;1  222  DMWSVGCIFA ELLHGKPIFP GKDEPEQLNK IFELCGAPDE VNWPGVTKTP
     CDK9Dm   243  DMWGAGCIMA EMWTRSPIMQ GNTEQQQLTF ISQLCGSFTP DVWPGVEELE
     CDK9Hs   211  DLWGAGCIMA EMWTRSPIMQ GNTEQHQLAL ISQLCGSITP EVWPNVDNYE
     CDK9Ce   190  DVWSTGCMLG ELFTRKPLFN GNNEFGQLEL ISKVCGSPNV DNWPELTELV

                               **                 *     * *    **    **
Arath;CDKC;1  273  WYNQMKSSRP LKRRVREIYR .H.FDRHALE LLEKMLVLDP SQRICAKDAL
Arath;CDKC;2  273  WHNNFKPARP LKRRVREFFR .H.FDRHALE LLEKMLVLDP AQRISAKDAL
Medsa;CDKC;1  272  WYNQFKPSRP MKRRLREVFR .H.FDRHALE LLEKMLTLDP AQRIPAKDAL
     CDK9Dm   293  LYKSIELPKN QKRRVKERLR PYVKDQTGCD LLDKLLTLDP KKRIDADTAL
     CDK9Hs   261  LYEKLELVKG QKRKVKDRLK AYVRDPYALD LIDKLLVLDP AQRIDSDDAL
     CDK9Ce   240  GWNTFRMKRT YQRRIREEFE .HIMPREAVD LLDKMLTLNP EKRISAKEAL
```

Figure 1

```
               ***           **** *       *   *         *       * ***   *
 cycT1At   34  WYFSREEIER FSPSRKDGID LVKESFLRSS YCTFLQRLGM KLHVSQVTIS
 cycT1Hs   11  WYFTREQLEN .SPSRRFGVD PDKELSYRQQ AANLLQDMGQ RLNVSQLTIN
cycT1Mou   11  WYFTREQLEN .SPSRRFGVD SDKELSYRQQ AANLLQDVGQ RLNVSQLTIN
  cycTDm   46  WYFSNDQLAN .LPSRRCGIK GDDELQYRQM TAYLIQEMGQ RLQVSQLCIN

                * *  ****     * *               ** * * ** *   *  *
 cycT1At   84  CAMVMCHRFY MRQSHAKNDW QTIATSSLFL ACKAEDEPCQ LSSVVVASYE
 cycT1Hs   61  TAIVYMHRFY MIQSFTQFPG NSVAPAALFL AAKVEEQPKK LEHVIKVAHT
cycT1Mou   61  TAIVYMHRFY MIQSFTQFHR YSMAPAALFL AAKVEEQPKK LEHVIKVAHT
  cycTDm   96  TAIVYMHRFY AFHSFTHFHR NSMASASLFL AAKVEEQPRK LEHVIRAANK

                                   *             *    * *   *    *
 cycT1At  134  IIYEWDPSAS IRIHQTECYH EFKEIILSGE SLLLSTSAFH LDIELPYKPL
 cycT1Hs  111  CLH...PQES LPDTRSEAYL QQVQDLVILE SIILQTLGFE LTIDHPHTHV
cycT1Mou  111  CLH...PQES LPDTRSEAYL QQVQDLVILE SIILQTLGFE LTIDHPHTHV
  cycTDm  146  C......... LPPTTEQNYA ELAQELVFNE NVLLQTLGFD VAIDHPHTHV

                       *    ***               *     ***     *       *
 cycT1At  184  AAALNRLNAW PDLATAAWNF VHDWIR.TTL CLQYKPHVIA TATVHLA
 cycT1Hs  158  VKCTQLVRAS KDLAQTSYFM ATNSLHLTTF SLQYTPPVVA CVCIHLA
cycT1Mou  158  VKCTQLVRAS KDLAQTSYFM ATNSLHLTTF SLQYTPPVVA CVCIHLA
  cycTDm  187  VRTCQLVKAC KDLAQTSYFL ASNSLHLTSM CLQYRPTVVA CFCIYLA
```

Figure 2

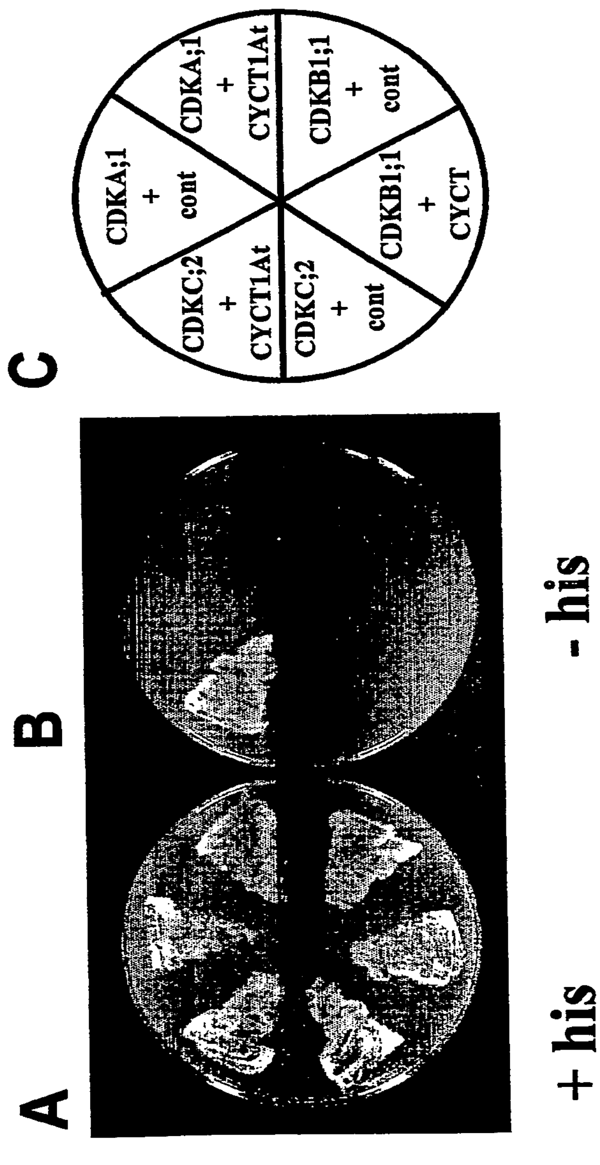
A
B
C
CDKA;1 + cont
CDKA;1 + CYCT1At
CDKB1;1 + cont
CDKB1;1 + CYCT
CDKC;2 + cont
CDKC;2 + CYCT1At
+ his
- his

Figure-3

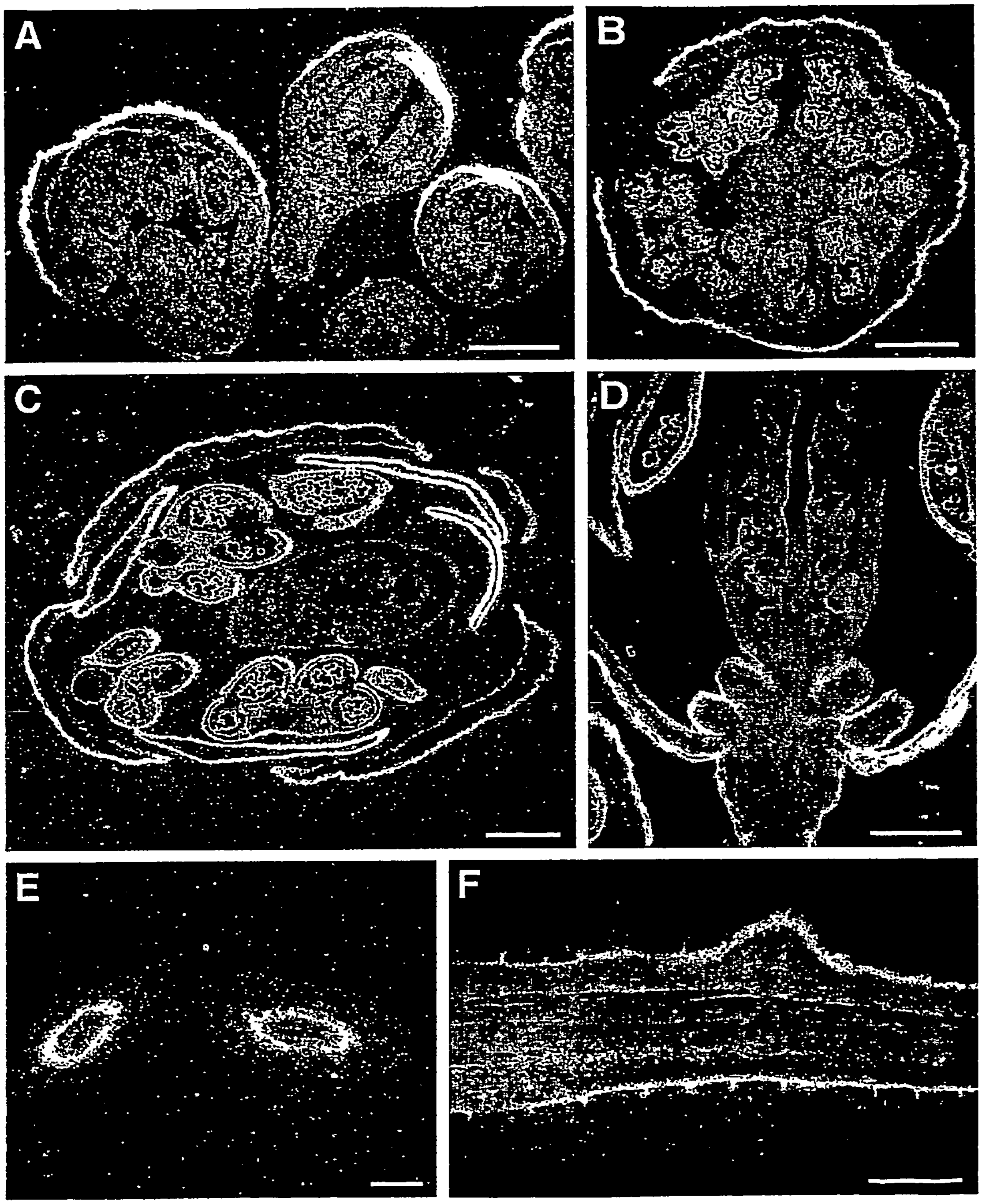

Figure 4

Seq ID info

Seq ID No	length	NA or AA	Start/stopcodon	Description
1	1738	NA	start on 91; stop on 1606	Arath;CDKC;2 (complete cDNA)
2	505	AA		Arath;CDKC;2 (complete protein)
3	954	NA	start on 1; stop on 952	CYCT1At (complete cDNA)
4	317	AA		CYCT1At (complete protein)
5	525	NA	start on 98; no stop	rubisco activase (partial cDNA)
6	142	AA		rubisco activase (partial protein)
7	657	NA	no start; no stop	DAG (partial cDNA)
8	219	AA		DAG (partial protein)
9	395	AA		DAG (complete protein)
10	639	NA	no start; no stop	RNP-like (partial cDNA)
11	213	AA		RNP-like (partial protein)
12	589	NA	no start; stop at 379	CDKCIP1 (partial cDNA)
13	126	AA		CDKCIP1 (partial protein)
14	1332	AA		CDKCIP1 (complete protein)
15	664	NA	start at 24; no stop	AtGT1 (partial cDNA)
16	213	AA		AtGT1 (partial protein)
17	406	AA		AtGT1 (complete protein)

Figure 5 : Nucleic acids and polypeptides of the invention

SEQ ID list

SEQ ID NO 1

TCTGAGAGAAAAGGAAAGCGATCGAGAAAGACGTAATTTGATCATCGGAG
TAAAAGATATTGTTCGACAGTGGGACTCCCGGGAACGAAGATGGCGATGG
CATCATTCGGGCAGCTAAATCTCGAGGAACCTCCTCCAATCTGGGGATCT
CGCAGCGTTGATTGCTTTGAGAAGCTCGAACAAATTGGTGAAGGCACTTA
CGGTCAAGTTTACATGGCTAAAGAAATCAAAACTGGTGAAATTGTGGCTC
TCAAAAAGATACGTATGGACAATGAAAGAGAAGGGTTTCCTATAACAGCT
ATTAGAGAGATAAAGATTCTGAAGAAGCTTCATCATGAAAATGTCATTCA
GCTGAAAGAGATTGTGACTTCACCAGGTCGGGACAGGGATGACCAAGGAA
AGCCAGATAATAACAAATACAAGGGTGGCATCTACATGGTTTTTGAGTAC
ATGGATCATGATTTGACTGGACTAGCTGATCGTCCTGGACTGAGATTTAC
TGTTCCTCAAATTAAGTGTTACATGAAGCAATTGCTTACCGGGCTTCACT
ATTGTCATGTGAATCAAGTGCTTCACCGTGATATAAAAGGCTCAAATCTC
CTTATCGACAATGAGGGAAATTTAAAGCTGGCTGATTTTGGGCTTGCACG
GTCGTATTCTCATGATCATACTGGAAATCTTACAAATCGTGTCATCACAT
TGTGGTATAGGCCCCCTGAATTACTACTTGGGGCTACAAAATATGGCCCA
GCAATTGACATGTGGTCGGTTGGTTGCATATTTGCCGAACTTTTGCATGC
AAAACCAATCTTACCTGGGAAAAATGAGCAAGAACAATTGAACAAGATAT
TTGAGCTTTGTGGATCACCTGATGAAAAACTTTGGCCTGGGGTTTCCAAG
ATGCCTTGGTTCAACAATTTCAAGCCTGCACGGCCCTTGAAGAGGCGTGT
AAGAGAGTTCTTCAGACACTTTGATCGGCATGCTCTTGAATTACTGGAGA
AAATGTTGGTGCTTGATCCAGCACAGAGAATATCGGCAAAGGATGCTCTT
GATGCCGAGTACTTTTGGACTGATCCGTTGCCATGTGACCCAAAGAGTCT
GCCCACATATGAATCATCACATGAGTTCCAGACAAAGAAAAAGCGGCAAC
AGCAGCGCCAAAACGAGGAAGCAGCAAAAAGACAGAAACTGCAGCATCCA
CCGCTGCAGCACTCTCGCTTACCCCCATTACAACATGGTGGACAGTCTCA
TGCTGCTCCACATTGGCCTGCAGGTCCAAACCATCCCACTAACAACGCAC
CACCACAAGTACCTGCGGGACCCAGCCACAACTTCTATGGAAGCCGCGT
GGTCCACCTGGTCCAAACCGCTACCCTCCTAGCGGAAACCAGAGCGGGGG
TTATAATCAAAGCCGAGGAGGTTACAGCAGTGGATCATATCCTCCACAAG
GACGTGGAGCTCCTTATGTGGCTGGTCCTAGAGGGCCTAGTGGTGGCCCG
TACGGGGTTGGACCTCCTAACTACACACAAGGTGGTCAGTATGGTGGCTC
TGGTAGCTCGGGAAGAGGGCAGAATCAGAGAAACCAGCAATACGGATGGC
AACAGTAAAGAGCTCTAATGATTTGTTGATCTGATATCTTACTCTTACAT
TTAATGTGATACTAGTAATAAGCTAATAATAGATTAATATGAGAAACTGA
ATCTCTTTCTTTTCCCAAAAAAAAAAAAAAAAAAAAAA

Figure 5 (continued) : Nucleic acids and polypeptides of the invention

SEQ ID NO 2

MAMASFGQLNLEEPPPIWGSRSVDCFEKLEQIGEGTYGQVYMAKEIKTGE
IVALKKIRMDNEREGFPITAIREIKILKKLHHENVIQLKEIVTSPGRDRD
DQGKPDNNKYKGGIYMVFEYMDHDLTGLADRPGLRFTVPQIKCYMKQLLT
GLHYCHVNQVLHRDIKGSNLLIDNEGNLKLADFGLARSYSHDHTGNLTNR
VITLWYRPPELLLGATKYGPAIDMWSVGCIFAELLHAKPILPGKNEQEQL
NKIFELCGSPDEKLWPGVSKMPWFNNFKPARPLKRRVREFFRHFDRHALE
LLEKMLVLDPAQRISAKDALDAEYFWTDPLPCDPKSLPTYESSHEFQTKK
KRQQQRQNEEAAKRQKLQHPPLQHSRLPPLQHGGQSHAAPHWPAGPNHPT
NNAPPQVPAGPSHNFYGKPRGPPGPNRYPPSGNQSGGYNQSRGGYSSGSY
PPQGRGAPYVAGPRGPSGGPYGVGPPNYTQGGQYGGSGSSGRGQNQRNQQ
YGWQQ

SEQ ID NO 3

ATGGGAGAGGAGCATCCGAGAAAGCGGTCTAGACAACATTTTGAAGCGGA
GGCGAGAAACGTATCGTTGTTTGAATCCCCTCAATGCGAAACCTCCAAGT
GGTATTTCAGCAGGGAAGAGATTGAGCGTTTCTCTCCATCCAGAAAAGAT
GGGATTGATCTTGTGAAGGAGTCGTTTTTACGGTCTTCGTATTGCACCTT
CCTGCAAAGACTTGGCATGAAGCTTCATGTGTCCCAGGTTACAATATCAT
GTGCAATGGTGATGTGCCACAGGTTTTACATGCGCCAATCTCATGCAAAA
AATGACTGGCAGACAATAGCGACTTCCAGTCTGTTCCTCGCTTGCAAAGC
TGAAGATGAGCCATGTCAACTGTCCAGTGTCGTTGTAGCGTCTTATGAAA
TAATTTATGAGTGGGATCCTTCTGCCTCAATTAGAATCCATCAAACTGAG
TGTTATCATGAATTTAAAGAAATTATTTTGTCCGGGGAAAGTCTTCTGCT
GAGCACAAGTGCTTTCCATTTAGACATTGAACTTCCCTACAAACCTCTGG
CTGCGGCTTTGAATAGACTGAACGCTTGGCCTGACCTTGCAACAGCTGCA
TGGAATTTTGTGCATGACTGGATTCGAACCACACTATGCTTGCAGTACAA
ACCCCATGTTATTGCAACAGCCACTGTGCACCTAGCTGCTACGTTTCAGA
ATGCGAAAGTAGGCAGCAGGAGAGATTGGTGGTTGGAGTTTGGAGTTACA
ACTAAGCTATTAAAAGAGGTAATCCAGGAGATGTGCACACTGATAGAAGT
GGACAGAAGGAGGAATATGCCACCTCCACTTCCACCTCCAAGAAGAGAGT
TAAGTTGGGCAATACCTGCAGCCGTAAAGCCGGTCCATATGGCTAGAGCT
TATCCGTTTCACAGCTACCCTTTGCAGTCCTATAGACAGGCTGGCATCTG
GTGA

Figure 5 (continued) : Nucleic acids and polypeptides of the invention

SEQ ID NO 4

MGEEHPRKRSRQHFEAEARNVSLFESPQCETSKWYFSREEIERFSPSRKD
GIDLVKESFLRSSYCTFLQRLGMKLHVSQVTISCAMVMCHRFYMRQSHAK
NDWQTIATSSLFLACKAEDEPCQLSSVVVASYEIIYEWDPSASIRIHQTE
CYHEFKEIILSGESLLLSTSAFHLDIELPYKPLAAALNRLNAWPDLATAA
WNFVHDWIRTTLCLQYKPHVIATATVHLAATFQNAKVGSRRDWWLEFGVT
TKLLKEVIQEMCTLIEVDRRRNMPPPLPPPRRELSWAIPAAVKPVHMARA
YPFHSYPLQSYRQAGIW

SEQ ID NO 5

GCTCACACCAACATCTCTCTAAGCTTCTTCTTCTACCAATCTAATTCCTC
TCTTCAGCTTCTTGTGTTGTGACGCATACTCGTCGCAGTCTTGAGATATG
GCCGCCGCAGTTTCCACCGTCGGTGCCATCAACAGAGCTCCGTTGAGCTT
GAACGGGTCAGGATCAGGAGCTGTATCAGCCCCAGCTTCAACCTTCTTGG
GAAAGAAAGTTGTAACTGTGTCGAGATTCGCACAGAGCAACAAGAAGAGC
AACGGGATCATTCAAGGTGTTGGCTGTGAAAGAAGACAAACAAACCGATG
GAGACAGAATGGAGAGGTCTTGCCTACGACACTTCTGATGATCAACAAGA
CATCACCAGAGGCAAGGGTATGGTTGACTCTGTCTTCCAAGCTCCTATGG
GAACCGGACTACCACGCTGTCCTTAGCTCATACGAATACGTTAGCCAAGG
CCTTAGGCAGTACAACTTGGACAACATGATGGATGGGTTTTACATTGCTC
CTGCTTTATGGACAAGCTTGTTGTT

SEQ ID NO 6

MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTFLGKKVVTVSRFAQSNKK
SNGIIQGVGCERRQTNRWRQNGEVLPTTLLMINKTSPEARVWLTLSSKLL
WEPDYHAVLSSYEYVSQGLRQYNLDNMMDGFYIAPALWTSLL

SEQ ID NO 7

GGCACGAGGCGATCCTTTGCTTCATCTGCTCCTCTCGCCAAATCTCCGGC
GTCTTCTCTGCTCAGCCGGTCCCGTCCTCTGGTCGCCGCTTTTTCCTCCG
TTTTCCGTGGCGGCCTTGTGTCTGTCAAAGGTCTCTCGACGCAGGCTACA
TCGTCTTCTCTGAACGACCCGAATCCCAACTGGTCGAACAGGCCACCTAA
GGAGACGATCCTGCTCGATGGTTGCGATTTCGAGCACTGGCTTGTCGTCG
TGGAGCCACCTCAGGGTGAGCCTACTAGAGATGAAATCATTGATAGCTAC
ATCAAAACCCTAGCTCAGATTGTTGGCAGTGAAGACGAAGCTAGGATGAA
GATCTACTCGGTTTCAACTAGGTGCTACTATGCTTTTGGGGCACTTGTGT

Figure 5 (continued) : Nucleic acids and polypeptides of the invention

CAGAAGATCTTTCTCACAAGTTAAAAGAGTTGTCAAATGTGCGCTGGGTT
CTTCCTGACTCTTACCTGGATGTGAGGAACAAAGACTATGGAGGGGAACC
TTTCATCGATGGGAAGGCTGTTCCTTATGATCCTAAGTACCACGAGGAGT
GGATAAGGAACAATGCTAGAGCAAATGAAAGGAACAGACGCAATGACCGT
CCTCGCAACAATGATAGAAGCAGAAACTTTGAGAGGAGAAGAGAGAACAT
GGCAGGA

SEQ ID NO 8

GTRRSFASSAPLAKSPASSLLSRSRPLVAAFSSVFRGGLVSVKGLSTQAT
SSSLNDPNPNWSNRPPKETILLDGCDFEHWLVVVEPPQGEPTRDEIIDSY
IKTLAQIVGSEDEARMKIYSVSTRCYYAFGALVSEDLSHKLKELSNVRWV
LPDSYLDVRNKDYGGEPFIDGKAVPYDPKYHEEWIRNNARANERNRRNDR
PRNNDRSRNFERRRENMAG

SEQ ID NO 9

MATHTISRSILCRPAKSLSFLFTRSFASSAPLAKSPASSLLSRSRPLVAA
FSSVFRGGLVSVKGLSTQATSSSLNDPNPNWSNRPPKETILLDGCDFEHW
LVVVEPPQGEPTRDEIIDSYIKTLAQIVGSEDEARMKIYSVSTRCYYAFG
ALVSEDLSHKLKELSNVRWVLPDSYLDVRNKDYGGEPFIDGKAVPYDPKY
HEEWIRNNARANERNRRNDRPRNNDRSRNFERRRENMAGGPPPQRPPMGG
PPPPPHIGGSAPPPPHMGGSAPPPPHMGQNYGPPPPNNMGGPRHPPPYGA
PPQNNMGGPRPPQNYGGTPPPNYGGAPPANNMGGAPPPNYGGGPPPQYGA
VPPPQYGGAPPQNNNYQQQGSGMQQPQYQNNYPPNRDGSGNPYQG

SEQ ID NO 10

GGCACGAGGCCACCTACTTTGACTGATGAAGAGTTTCGCCAGTACTTTGA
AGTTTATGGCCCTGTGACTGATGTTGCAATCATGTATGACCAGGCTACCA
ACCGTCCTCGTGGGTTTGGATTTGTTTCCTTCGACTCTGAAGATGCGGTA
GACAGTGTTTTGCACAAGACTTTCCATGATTTGAGCGGTAAACAAGTTGA
AGTAAAGCGTGCTCTTCCTAAAGATGCCAATCCTGGAGGTGGTGGACGAT
CAATGGGTGGTGGTGGCTCTGGTGGTTACCAGGGTTATGGTGGCAATGAA
AGCAGTTATGATGGACGTATGGATTCCAATAGGTTTTTGCAGCATCAAAG
TGTTGGAAATGGTTTACCATCTTATGGTTCTTCTGGTTATGGCGCTGGCT
ATGGAAATGGTAGTAATGGTGCCGGGTATGGTGCCTATGGAGGTTACACT
GGTTCTGCTGGAGGTTATGGCGCTGGTGCTACTGCTGGATATGGAGCAAC
GAACATTCCAGGTGCTGGCTATGGAAGTAGTACTGGAGTTGCTCCGAGAA
ACTCATGGGACACTCCAGCTTCTAATGGTTATGGGAACCCAGGCTATGGG
AGTGGTGCTGCTCATAGTGGATATGGAGTTCCTGGTGCA

Figure 5 (continued) : Nucleic acids and polypeptides of the invention

SEQ ID NO 11

GTRPPTLTDEEFRQYFEVYGPVTDVAIMYDQATNRPRGFGFVSFDSEDAV
DSVLHKTFHDLSGKQVEVKRALPKDANPGGGGRSMGGGGSGGYQGYGGNE
SSYDGRMDSNRFLQHQSVGNGLPSYGSSGYGAGYGNGSNGAGYGAYGGYT
GSAGGYGAGATAGYGATNIPGAGYGSSTGVAPRNSWDTPASNGYGNPGYG
SGAAHSGYGVPGA

SEQ ID NO 12

AAATGGTCCTACGAGATGTTTGATTTCCGAGGATATTCAGACCTTGATAG
AGGGATTGGTCAAAGAGAATATCCACAGCAGTACGGTGGGCACTTGGACC
CCATGTTAGCTCCTCCTCCTCCAAATCTGATGGACAATGCATTCCCA
TTGCAACAACGTTATGCGCCTCATTTCGATCAAATGAATTACCAGAGGAT
GAGCTCTTTCCCACCTCAGCCTCCATTGCAACCTAGCGGACATAATCTCT
TAAATCCTCATGACTTTCCACTGCCACCGCCACCACCTAGTGACTTCGAA
ATGAGTCCAAGGGGTTTTGCCCCTGGCCCGAACCCGAACTACCCTTATAT
GAGTCGATCTGGCGGTTGGATTAATGACTAGATCAGCACTCATTATCCTT
GTAGTTGCAACATTAGTAGTTTGATTGATCTTTTGTGTCTCACTCTACGA
AAGTGTAGGAAGAATAGAAGAAATCTATAACTTTTCTCTGCCACTCACAT
GTGTAGCTAGTGGGCCTTTTAGCTGTTTAATAATATAAAAGAAAAAGAAG
CCAGCTTCTATTGTCTTAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO 13

KWSYEMFDFRGYSDLDRGIGQREYPQQYGGHLDPMLAPPPPPNLMDNAFP
LQQRYAPHFDQMNYQRMSSFPPQPPLQPSGHNLLNPHDFPLPPPPPSDFE
MSPRGFAPGPNPNYPYMSRSGGWIND

SEQ ID NO 14

MTFVDDDEEEDFSVPQSASNYYFEDDDKEPVSFARLPIQWSVEEKVDGSG
LGFYLRGRSDNGLLPLHKLVKAWRYDLSNFQPEISVLTKDNIWIKLEEPR
KSYGELIRTVLVTLHSIQFLRRNPQASEKALWEKLTRSLRSYDVKPSQND
LVDHIGLIAEAAKRDRNLANSKFILAFLTKKPTKRRLPDEDNAKDDFIVG
DEDTYVASDEDELDDEDDDFFESVCAICDNGGEILCCEGSCLRSFHATKK
DGEDSLCDSLGFNKMQVEAIQKYFCPNCEHKIHQCFICKNLGSSDNSSGA
AEVFQCVSATCGYFYHPHCVTRRLRLGNKEESEALERQIIAGEYTCPLHK
CSVCENGEVKTDSNLQFAVCRRCPKSYHRKCLPREISFEDIEDEDILTRA

Figure 5 (continued) : Nucleic acids and polypeptides of the invention

WDGLLHNRVLIYCQEHEIDEELLTPVRDHVKFPFTEEQKVFVKEQRRILE
SHVGRDKARLKVKDPALQDTCGKASKNSFRSSFPSSKDGFSTKKHGLVSS
VPDHSRKRKDIDPSIKHKMVPQKSQKMMEDSREAGKNKLGVKEARDAGKS
KISLGERLFSYTQEPNPVKPGRVIPVDSKHNKTDSIASKEPGSEIPTLDN
DSQRRLLAVMKKATEEITMGTILKKFKIQSTMSTHSTRNVVDKTITMGKV
EGSVQAIRTALKKLEEGGNIEDAKAVCEPEVLSQILKWKDKLKVYLAPFL
HGARYTSFGRHFTNPEKLQQIVDRLHWYADDGDMIVDFCCGSNDFSCLMN
AKLEETGKKCLYKNYDLFPAKNNFNFERKDWMTVSKDELEPGSKLIMGLN
PPFGVNASLANKFITKALEFRPKILILIVPPETERFQFPSISSAPLYHSI
TLIYRLLSLSLVKSITFLNRLDKKKSSYVLIWEDKTFLSGNSFYLPGSVN
EEDKQLEDWNLVPPPLSLWSRSDFAAKHKKIAEKHCHLSRDVGSSKLKIV
EEEANASLHPLGASDGMCDDIPMEKDELEVAECVNKILVSEKIDTVETVA
RVHQSDHLSRRSQLKKEGKTKDYSGRKLGKSMDSNNVDWKSNDMEEDQGE
LSRAPESIKVKIPEMTSDWQSPVRSSPDDIYAVCTSISTTTPQRSHEAVE
ASLPAITRTKSNLGKNIREHGCKVQGTGKPEVSRDRPSSVRTSREDIYTV
RPSPENTGQKPFEAFEPSYGASLSHFDDGLAAKYGGFGGGYRMPDPPFLP
DQFPLRNGPNEMFDFRGYSDLDRGIGQREYPQQYGGHLDPMLAPPPPPNL
MDNAFPLQQRYAPHFDQMNYQRMSSFPPQPPLQPSGHNLLNPHDFPLPPP
PPSDFEMSPRGFAPGPNPNYPYMSRSGGWIND

SEQ ID NO 15

GGCACGAGCCGGGAAACACAGCAATGTTCATTTCCGACAAATCTCGTCCT
ACTGATTTCTACAAAGACGATCATCACAATTCCTCCACCACCAGCACCAC
ACGCGATATGATGATCGATGTACTCACCACTACCAACGAATCAGTAGATC
TACAATCTCACCACCACCACAATCACCACAATCATCATCTCCACCAATCT
CAGCCACAACAACAGATTCTCCTCGGAGAAAGCAGTGGAGAAGATCACGA
AGTTAAAGCACCAAAGAAACGAGCGGAGACATGGGTTCAAGACGAAACTC
GTAGCTTAATCATGTTCCGTAGAGGTATGGATGGTTTATTCAATACATCC
AAATCTAATAAACATCTCTGGGAACAGATTTCGTCTAAGATGAGAGAAAA
AGGGTTTGATCGATCTCCGACTATGTGTACTGATAAATGGAGGAATCTGT
TGAAAGAGTTTAAGAAAGCTAAGCATCATGATAGAGGAAATGGATCGGCG
AAGATGTCGTATTACAAAGAGATTGAAGATATTCTTAGAGAGAGGAGCAA
AAAAGTGACACCACCACAGTATAATAAGAGCCCTAATACACCACCTACAT
CAGCCAAAGTTGATTCCTTTATGCAATTTACTGATAAAGGGTTTGATGAT
ACGAGCATTTCTTT

Figure 5 (continued) : Nucleic acids and polypeptides of the invention

SEQ ID NO 16

MFISDKSRPTDFYKDDHHNSSTTSTTRDMMIDVLTTTNESVDLQSHHHHN
HHNHHLHQSQPQQQILLGESSGEDHEVKAPKKRAETWVQDETRSLIMFRR
GMDGLFNTSKSNKHLWEQISSKMREKGFDRSPTMCTDKWRNLLKEFKKAK
HHDRGNGSAKMSYYKEIEDILRERSKKVTPPQYNKSPNTPPTSAKVDSFM
QFTDKGFDDTSIS

SEQ ID NO 17

MFISDKSRPTDFYKDDHHNSSTTSTTRDMMIDVLTTTNESVDLQSHHHHN
HHNHHLHQSQPQQQILLGESSGEDHEVKAPKKRAETWVQDETRSLIMFRR
GMDGLFNTSKSNKHLWEQISSKMREKGFDRSPTMCTDKWRNLLKEFKKAK
HHDRGNGSAKMSYYKEIEDILRERSKKVTPPQYNKSPNTPPTSAKVDSFM
QFTDKGFDDTSISFGSVEANGRPALNLERRLDHDGHPLAITTAVDAVAAN
GVTPWNWRETPGNGDDSHGQPFGGRVITVKFGDYTRRIGVDGSAEAIKEV
IRSAFGLRTRRAFWLEDEDQIIRCLDRDMPLGNYLLRLDDGLAIRVCHYD
ESNQLPVHSEEKIFYTEEDYREFLARQGWSSLQVDGFRNIENMDDLQPGA
VYRGVR

Figure 5 (continued) : Nucleic acids and polypeptides of the invention

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.